22016-1227/318869

IN THE UNITED STATES PATENT OFFICE

In re patent application of:

Albert S. Feng et al.

Application No. 09/193,058

Filed: November 16, 1998

BINAURAL SIGNAL PROCESSING TECHNIQUES

Before the Examiner:
Andrew R. Graham

Group Art Unit: 2644

DECLARATION UNDER 37 CFR §1.132

I, Douglas L. Jones, the Declarant, hereby declare as follows:

1. I am currently a Professor of Electrical and Computer Engineering engaged in research in the area of acoustic signal processing and hearing aids.

2. I have expertise in the fields of electronics, signal processing, acoustics, and hearing aids. My work experience, education, and other credentials in these fields are further documented as attached hereto in Exhibit A.

3. I have reviewed U.S. Patent No. 09/193,058 filed November 16, 1998 (the "Subject Application"), and pending claims that have been or will be proposed as attached hereto in Exhibit B. (Collectively designated the "Patent Claims" hereinafter).

4. I have also reviewed the United States Patent Office Action mailed 27 September 2004

(hereinafter the "Office Action") that rejects the Subject Application, and I have reviewed the references asserted in the Office Action, U.S. Patent No. 6,002,776 to Bhadkamkar, U.S. Patent No. 4,601,025 to Lea, and U.S. Patent No. 5,581;620 to Brandstein et al., all of which are attached in Exhibit C.

5. The documents of Exhibits B and C are all directed to technologies in which I have expertise.

6. Based on my review of Exhibits B and C, those skilled in the art would be discouraged from considering the combination of Bhadkamkar, Lea, and Brandstein as proposed in the Office Action. The operability and/or suitability for intended use would be undermined by the asserted combinations. Correspondingly, those skilled in the art would not reasonably expect success to result. For instance, the Office Action proposes to "incorporate an individual correlator system as taught by Lea as part of the DOA estimator of the system of Bhadkamkar." (Office Action, page 5). The Office Action indicates that an "individual correlator system" includes at least delay lines 42 and 44 and correlators 46. Correlators 46 each provide an output, three of which are specifically designated by d, e, and f. If this is the interface to separator 30 intended in the Office Action, using only some of the outputs or combining them in some manner would likely undermine operability of the proposed combination -- or at the very least require undue experimentation and/or significant undisclosed modifications to succeed.

7. In another instance, the Office Action cites to column 4, lines 59-66 of Lea stating that

the "identification of a peak indication gives rise to the time differential between the angle of incidence of the received sound and the phase center of the two receiving elements." Based on a review of column 4, line 59 - column 5, line 10 of Lea, it does not appear to support this statement. It is surmised that the Office Action could be contending that the referenced time differentials of Lea are being interfaced with the separator 30 of Bhadkamkar -- that is t_1 and/or t_2 rather than the correlator outputs (d, e, f) are provided to separator 30. However, as perhaps best shown in Fig. 4b of Lea, the derivation of t_1 and of t_2 depends on the inputs a, b, and c from two other delay lines 51 and 52, and associated correlators 53 in addition to inputs d, e, and f from correlators 46. Correspondingly, the inclusion of this additional hardware (four delay lines and numerous correlators 46 and 53) is inconsistent with the two input, two output DOA estimator 20 of Bhadkamkar. Moreover, it is questionable that separator 30 would properly work with t_1 and/or t_2 as defined by Lea. Indeed, Lea depends on subtraction network 54 to provide a signal difference, which is needed to subsequently determine incidence angle with computer 55. Even with further speculative, undisclosed modifications, the prospect of success remains dubious.

8. Based on my review of Exhibits B and C, there does not appear to be any teaching, suggestion, or disclosure of generation of a characteristic signal representative of the desired acoustic signal that occurs during performance of determining location of the second source; where the interfering signal is from this second source.

9. Attached as Exhibit D is a copy of a publication in the Journal of the Acoustical Society

of America (hereinafter "JASA"). JASA is one of the pre-eminent scholarly publications in the field to which the Patent Claims pertain. Specifically, "A two-microphone dual delay-line approach for extraction of a speed sound in the presence of multiple interferers", JASA 110(6) (December 2001) (hereinafter the "JASA Article") describes the virtues of the inventions defined by at least claims 34, 46, 61, and 62 in connection with section II.A. Furthermore, the discussion in section I. of the JASA Article describes the long-standing desire for a way to extract signals subject to the "cocktail party" effect. As described in the Subject Application, page 14, line 24 – page 15, line 32, the inventions defined by claims 34-35, 37-42, 44-48, 50-54, and 56-66 provide significant advances over the state of the art by providing source separation of as little as 2 degrees arithmetically and extracting a desired signal of lesser intensity than a close interferer to address this long-felt desire.

10. Based on information and belief, the owner of all right, title and interest to the Subject Application, is the Board of Trustees of the University of Illinois (hereinafter the "Owner"). The Owner has licensed the subject patent application to Phonak AG by written agreement dated November 1, 2000.

11. The undersigned, being hereby warned that willful false statements and the like are punishable by fine or imprisonment, or both (18 U. S C. 1001), and may jeopardize the validity of the application or any patent issuing thereon, declares that all statements made of his/her own knowledge are true and that all statements made on information and belief are believed to be true.

Douglas L. Jones
[Declarant Name]

January 21, 2005
Date

Douglas L. Jones

OFFICE ADDRESS

Coordinated Science Laboratory
1308 W. Main Street
University of Illinois
Urbana, IL 61801
(217) 244-6823
(217) 244-1642 (FAX)
dl-jones@uiuc.edu

HOME ADDRESS

1214 W. Church St.
Champaign, IL 61821
(217) 359–2038

EDUCATION

Rice University Ph.D. in Electrical and Computer Engineering, 1987
MSEE in Electrical and Computer Engineering, 1985
BSEE in Electrical Engineering (Summa Cum Laude), 1983

AWARDS and HONORS

Fellow, IEEE, 2002
Connexions Author of the Year, 2003
Fulbright Junior Research Fellowship, 1987-1988
National Science Foundation Graduate Fellowship, 1983-1986
Member of Phi Beta Kappa, Tau Beta Pi, and Eta Kappa Nu Honor Societies

EXPERIENCE

1988-Present **University of Illinois at Urbana-Champaign**
Professor, Dept. of Electrical and Computer Engineering, 1998-Present
Research Professor, Coordinated Science Laboratory
Research Professor, Beckman Institute
Associate Professor, 1993-1998
Assistant Professor, 1988-1993

Research and teaching in the area of signal and image processing.

Spring 2002 **University of California, Berkeley**
Visiting Scholar

Spring 1999	**Rice University** Texas Instruments Visiting Professor
Summer 1998	**University of Cambridge** Participant, Programme on Nonlinear and Nonstationary Signal Processing Isaac Newton Institute for Mathematical Sciences
Spring 1995	**University of Washington** Visiting Scientist
1987-1988	**Universität Erlangen-Nürnberg, Germany** Fulbright Postdoctoral Research Fellow
1984-Present	**Consultant** Consultant for several firms on the topics of magnetic field modeling, high-resolution NMR spectral analysis, optical time-domain reflectometry, sonar signal analysis, adaptive equalization, and electrocardiogram analysis.
1983-1987	**Rice University** Research Assistant, Dept. of Electrical and Computer Engineering Research on time-frequency and time-varying signal analysis, and FFT algorithms. Also developed a signal processing laboratory course and wrote an associated text-book, published by Prentice-Hall.

PROFESSIONAL ACTIVITIES

Member-at-large of the Board of Governors of the IEEE Signal Processing Society, 2002-2004

Co-Chairman	NSF/ONR Workshop on Signal Processing in Manufacturing and Machine Monitoring, March 1996, Alexandria, Virginia
Co-Chairman	Allerton Conference on Communications, Control, and Computing, October 2000, October 2001
Associate Editor	IEEE Signal Processing Letters, 1997-1999
Technical Program Comm.	1994 IEEE/SP Symp. on Time-Frequency and Time-Scale Analysis
Panels	NSF Research Initiation Award Panel, 1994 NSF Small Business Innovative Research Panel, 1995 NSF Signal Processing Systems Review Panel, 1998

Reviewer	*IEEE Transactions on Signal Processing* *IEEE Transactions on Information Theory* *IEEE Transactions on Biomedical Engineering* *IEEE Transactions on Aerospace and Electronic Systems* *IEEE Transactions on Circuits and Systems* *IEEE Signal Processing Letters* *IEEE Spectrum Magazine* *Proceedings of the IEEE* *Circuits, Systems, and Signal Processing* *Applied Signal Processing* IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing IEEE/SP Symp. on Time-Frequency and Time-Scale Analysis Allerton Conference on Communications and Control Midwest Symposium on Circuits and Systems National Science Foundation University of Illinois Campus Research Board McGraw-Hill Book Company Prentice-Hall Publishers, Inc.
Member	IEEE Signal Processing Society IEEE Communications Society

RECENT DEPARTMENTAL AND CAMPUS SERVICE

Department	ECE Curriculum Committee, 1990-1994, 1999-2001, 2002-Present Beckman Institute Program Advisory Committee, 2000-2001, 2002-present Secretary, Curriculum Revision Subcommittee, 1993-1994 Fellowship Committee, 1995-2001 Computer Resources Committee, 1994-1995 Chair, Computer Resources Committee, 1996-1997 Chair, Circuits and Signal Processing Area Committee, 1995-present CEPS Industrial Affiliates Program member, 1989-present CEPS Rockwell Corp. Liaison, 1993-present CEPS Steering Committee, 1993-1994, 1996-present ECE 320 Course Director, 1989-present
College	College of Engineering Executive Committee, 1997-2000 Associate Director, UI Motorola Center for Communications, 1997-present
Campus	University Senate, 1992-1994, 1996-1998, 1999-2001 Senate Educational Policy Committee, 1993-1994 Chair, Senate Subcommittee on Tuition Surcharge Policy, 1993-1994 Senate Council Coord. Group on Instructional Resource Improvement and Funding, 1994

Senate Academic Calendar Committee, 1996-1998

GRANTS AND CONTRACTS

NSF Research Initiation Award MIP-8908777, entitled "Area–Efficient VLSI Implementation of Signal Processing Algorithms via Multiple Coefficient Recoding," June 1989 through November 1991, amount: $63,119.

JSEP Contract N00014-89-C-0149 Unit 17, entitled "Adaptive Signal Processing," duration: August 1989 through July 1992, amount: $360,000, with Profs. H.V. Poor, W.K. Jenkins, and K.S. Arun.

NSF MIP 90-12747, entitled "Modeling of Time-Varying Signals," duration: 9/15/90–2/29/93, amount: $148,694, with Prof. K.S. Arun.

State of Illinois SCCA 91-82121, entitled "A Technology Transfer Initiative for ACT Signal Micro-Processor Technology," duration: 12/1/90–6/30/91, amount: $300,000, with Profs. W.K. Jenkins, D.L. Jones, D.C. Munson, Jr., Y. Bresler, N. Ahuja, and S. Franke.

DARPA N00014–91–J–1844, entitled "A Technology Transfer Initiative for ACT Signal Micro-Processor Technology," duration: 6/15/91–3/14/92, amount: $150,000, with Profs. W.K. Jenkins, D.L. Jones, D.C. Munson, Jr., Y. Bresler, N. Ahuja, and S. Franke.

JSEP Contract N00014-89-C-0149 Unit 19, entitled "Multi-Dimensional Adaptive Signal Processing," duration: October 1992 through September 1995, amount: $280,000, with Profs. W.K. Jenkins and M.T. Orchard.

Naval Surface Warfare Center, N00167-93-M-3228, entitled "Energy Partitioning in the Time-Frequency Plane," duration: March 1993 through February 1994, amount: $24,935.

Naval Surface Warfare Center, N000167-94-M-7133, entitled "Energy Partitioning in the Time-Frequency Plane," duration: March 1994 through February 1995, amount: $24,982.

DARPA, University of Minnesota subcontract entitled "Scalable Library for Digital Signal Processing," duration: March 1994 through December 1996, amount: $102,889, with Profs. K. Gallivan, D. Munson, and M. Orchard.

Center for Research on Applied Signal Processing, USC Subcontract PO#665433, entitled "Optimal and Adaptive Time-Frequency Methods for Detection and Estimation," duration: September 1994 through August 1996, amount: $90,000.

ONR AASERT Grant N00014-95-1-0907, entitled "Energy Partitioning Using Overdetermined Basis Decompositions," duration: May 1995 through May 1998, amount: $90,356.

ONR Contract N00014-95-1-0674, entitled "Adaptive and Optimal Time-Frequency Methods for Nonstationary Signals," duration: June 1995 through September 1998, amount: $210,000.

JSEP Contract N00014-96-1-0129 Unit 11, entitled "Acquisition and Demodulation for Wireless Communications," duration: October 1995 through September 1998, amount: $360,000, with Profs. D.V. Sarwate and U. Madhow.

Hewlett-Packard Equipment Gift No. 34748, entitled "Equipment Proposal for Communications Laboratories," duration: April 1997, amount: $71,980, with Profs. S. Franke and G. Papen.

NSF Grant no. MIP-9707742, entitled "Unified Algorithms and Architectures for Low-Power Wireless Video," duration: September 1997 through August 2001, amount: $479,878, with Profs. K. Ramchandran and N. Shanbhag.

ONR Contract N00014-95-1-0674 (extension), entitled "Time-Frequency-Space Processing And Multi-Component Signal Classification" duration: February 1998 through September 2001, amount: $182,994.

Texas Instruments DSP Equipment Gift for ECE DSP Laboratory, date: October 1998, amount: $17,955.

PhysioControls contract, entitled "Noise-Resistant Cardiac Arrhythmia Detector," duration: Jan 1998 through August 1999, amount: $23,325

University of Illinois Mary Jane Neer Research Fund, "High-Performance Dual Channel DSP-Based Acoustic Processor for Use in Hearing Aids," duration: January 1998 through December 1999, amount: with Profs. R. Bilger, A. Feng C. Lansing, W. O'Brien, and B. Wheeler

NIH, National Cancer Institute Grant no. CA079179, entitled "*In Vivo* Ultrasonic Microprobe for Tumor Diagnosis," duration: amount: $125,906 annual direct costs with Profs. W.D. O'Brien, Jr., J.F. Zachary, D.A. Payne, C. Liu

NSF contract CCR-9979381, entitled "An Integrated Exploration of Wireless Network Communication," duration: October 1999 through September 2001, amount: $466,636, with Profs. B. Hajek, R. Blahut, U. Madhow, and N. Shanbhag

CRASP, entitled "Modulation Classification," duration: October 1999 through September 2000, amount: $36,000

Texas Instruments DSP Equipment Gift for ECE DSP Laboratory, date: February 2000, amount: $42,985.

Motorola Corp. contract, entitled "Joint Source-Channel Matching for Wireless Multimedia Communication," duration: January 2000 through August 2002, amount: $180,000, with Prof. N. Shanbhag

NIH contract, entitled "Real Time Implementation of Intelligent Hearing Aids," duration: July 2000 through June 2002, amount: $254,804, with Profs. B. Wheeler, A. Feng, W.D. O'Brien, C. Lansing, and R. Bilger

DARPA contract, entitled "Development of a Biomimetic Acoustic Microsensor," duration: Sept. 29, 2000 through Sept. 28, 2003, amount: $3,000,000 (UIUC Subcontract $750,330), with Profs. R. Miles (SUNY Binghamton), A. Feng, B. Wheeler, W.D. O'Brien, and C. Liu

NSF contract, entitled "A Comprehensive Retargetable Embedded Systems Software Development Environment," duration: September 2000 through August 2005, amount: $4,000,000, with Profs. J. Davidson (UVa), D. Whalley (FSU), and K. Gallivan (FSU)

NSF contract CCR-00-85929 ITR, entitled "High-Speed Distributed Wireless Communication Networks," duration: September 2000 through August 2003, amount: $1,814,162, with Profs. B. Hajek, R. Blahut, R. Koetter, U. Madhow, S. Meyn, D. Sarwate, A. Singer, N. Shanbhag, and R. Srikant.

NSF contract , entitled "Integrated Sensitive Skin with Advanced Data Architecture," duration: June 2000 through May 2003, amount: $300,147, with Profs. C. Liu and N. Shanbhag

Phonak USA contract, entitled "Binaural Hearing Aids," duration: October 2000 through October 2003, amount: $1,210,000, with Profs. B. Wheeler, A. Feng, W.D. O'Brien, C. Lansing, and R. Bilger

NSF ITR contract CCR-0205638, entitled "ITR: Collaborative Hardware-Software Adaptation for Multimedia Applications," duration: October 1, 2002 through September 30, 2005, amount: $1,000,000, with Profs. S. Adve, R. Kravets, and K. Nahrstedt

NSF ITR contract CCR-0312432, entitled "ITR - Remote Reality: 4-D Audio-Visual Reconstruction and Compression from Multiple Sensors," duration: July 15, 2003 through June 30, 2006, amount: $326,735, with Prof. M. Do

NIH-NIDCD contract 1R01DC005782-01A1 (UIUC subaward 30568/1035968 from SUNY Binghamton), entitled "Sensing and Processing for Directional Hearing Aids," duration: October 1, 2003 through September 30, 2007, amount: (subcontract) $1,078,300, with Prof. R.N. Miles (SUNY Binghamton) and L. Degertekin (Georgia Tech)

PERSONAL

Born in Dallas, Texas on January 17, 1961. Married to Catherine A. Schmidt-Jones, three children.

MAJOR RESEARCH ACOMPLISHMENTS

Time-frequency and joint signal analysis:

- **Adaptive time-frequency representations.** Developed the concept of adaptive time-frequency representations (TFRs), and several of the leading signal-dependent and adaptive transforms. These include adaptive window short-time Fourier transforms, adaptive wavelet transforms, adaptive optimal kernels, adaptive cone kernels, and fast algorithms for their computation. Our recent "consistent time-frequency representation" achieves unprecedented resolution and cross-term suppression.
- **Statistical time-frequency analysis.** Developed a fundamental and comprehensive theory of statistical time-frequency analysis. We derived optimal kernels for nonstationary spectrum estimation and time-frequency estimation of TFRs of noisy or random signals. A theory of time-frequency detection has determined the class of detection problems for which time-frequency-based detection is globally optimal, the optimal kernels for these classes, and efficient methods of implementing these detectors. These techniques have been applied in a number of applications, including machine-fault detection and diagnosis, microembolus detection, and ECG classification.
- **Time-frequency-space processing.** Developed efficient, near-optimal time-frequency-space detectors and estimators for partially coherent arrays. Highly efficient, nearly optimal quadratic narrowband array detection algorithms have been an important by-product of this research.
- **Generalized joint signal representations.** Contributed several advances in generalized joint signal representations, including a time-frequency-based derivation of the chirplet transform, new orthogonal chirped bases, and unitarily transformed joint signal representations. We showed the equivalence of Cohen's and Baraniuk's methods for constructing general joint signal representations and contributed new insights to this theory. We extended the theory of adaptive and statistically optimal TFRs to generalized joint signal representations. We have developed four-parameter joint quadratic time-frequency-delay-doppler representations for applications such as improved adaptive time-frequency analysis and detection and classification.
- **Nonstationary blind source separation and interference cancellation.** Introduced new adaptive methods for blind source separation of nonstationary signals. These methods are simpler than existing methods and continuously track environmental changes. New methods for extraction of speech signals in cluttered environments (the "cocktail party" environment) have yielded great improvements over existing methods (see expanded discussion below).

Wavelet techniques and applications:

- **Denoising.** Innovative new methods for denoising multichannel data provide much better performance than single-channel methods and are very efficient; applications to hyperspectral imagery have shown more than 10 dB SNR gain. Derived worst-case bounds for the performance of denoising methods for both orthonormal bases and overdetermined frames. New "Bayesian pursuit" methods offer improve denoising performance for using overdetermined frames and a hierarchical statistical signal model.

- **Generalized wavelet decompositions and transforms.** Developed new orthogonal chirped wavelet bases, unitarily transformed basis decompositions, and efficient algorithms and implementations. Developed the chirplet transform from a time-frequency context.
- **MEMS sensors.** In a joint project with Profs. Chang Liu and Naresh Shanbhag, we are developing distributed, multi-element touch and flow sensors and embedded architectures and algorithms for extracting sophisticated touch and flow information, such as texture, turbulence, softness, and slippage to create an "artificial skin." In collaboration with Prof. Ron Miles at SUNY Binghamton, we are developing algorithms for high-accuracy direction-finding and signal recovery using colocated arrays of directional MEMS sensors.

Adaptive signal processing:

- **Blind equalization.** Developed a vector constant modulus algorithm for blind equalization of shaped channels, the first practical solution to this problem.
- **Nonlinear adaptive filters.** Developed a low-complexity, LMS-like algorithm for general systems with a memoryless nonlinearity. Application to nonlinear echo cancellation demonstrated substantial improvement over
- **Nonstationary blind source separation.** Introduced new adaptive methods for blind source separation of nonstationary signals, including both instantaneous and convolutive (dynamic) mixtures. These methods are simpler than existing methods and continuously track environmental changes. Research continues on faster algorithms.
- **Nonstationary adaptive beamforming.** Developed a frequency-domain minimum-variance distortionless-response beamformer for small arrays with unprecedented performance in the recovery of speech in nonstationary interference.
- **Algorithms.** Developed reduced-complexity and reduced-delay implementations for adaptive filters. Analyzed the transpose-form implementation for FIR adaptive filters and demonstrated its advantages for high-speed pipelined implementation.

Biomedical applications:

- **Binaural and directional hearing aids.** A multidisciplinary group based in the Beckman Institute has developed new binaural algorithms for the extraction of a desired source from a cluttered acoustic environment, such as in a restaurant or cocktail party. The new methods show remarkable performance improvements over conventional techniques in such environments, and have been implemented in a real-time DSP-based system. A new array technology using combinations of directional and omni microphones shows great promise for obtaining higher directivity in small (e.g., BTE) packages. Major research efforts toward commercialization for advanced hearing aids and other acoustic extraction applications such as hands-free telephony, automotive and military applications, and noise suppression for speech recognition systems continue.
- **Electrocardiogram analysis.** Developed improved methods for denoising electrocardiograms using adaptive time-frequency processing.
- **Microembolus detection.** Developed wavelet-based and chirped wavelet detectors from ultrasound reflections from microemboli. Theoretical and experimental studies demonstrated that these methods approach optimal performance.

- **Ultrasound image formation and analysis.** Developed fast frequency-domain three-dimensional reconstruc tion algorithms for image reconstruction from measurements on a circular aperture. Applications include high-resolution imaging from ultrasound micro-probes on the end of a needle and from small ultrasound catheters. Developed methods for detecting edges and tissue boundaries in ultrasound images.

  In collaboration with Prof. W.D. O'Brien, developed new methods for aberration correction in ultrasound imaging that perform much better than existing approaches in severe aberration.

- **Edge detection in ultrasound images.** Developed near-optimal, computationally efficient methods for detecting boundary segments in ultrasound speckle images.

- **fMRI Image Denoising.** Are developing (In collaboration with Prof. Farzad Kamalabadi and Dr. Keith Thulborn at UIC) efficient methods for blind removal of noise from functional MRI image sequences. These techniques will allow precise imaging at much faster rates by greatly reducing the necessary averaging time to construct low-noise functional images.

Telecommunications and other applications:

- **Peak Power Reduction for OFDM systems.** Developing new methods providing unprecedented peak-to-average power ratio (PAR) reduction for large-constellation OFDM systems. These methods are based on novel constellation-shaping approaches, and obtain these reductions with no loss of data rate or increase in symbol error rate. Waveform-modification methods that offer more modest reductions but are compatible with current standards have also been developed. Extension of these ideas to peak power reduction in optical and CDMA communication systems continues.

- **Optimal discrete multi-tone (DMT) power allocation.** Developed the first fast, exactly optimal algorithm for power allocation in discrete multi-tone modulation.

- **Instantaneous frequency estimation/FM demodulation.** Developed an adaptive TFR-based IF estimator for FM demodulation that lowers the SNR threshhold by 3-4 dB over existing methods.

- **Joint source-channel matching.** In collaboration with Profs. Shanbhag and Ramchandran, developed joint source-channel coding methods for wireless image and video transmission. These general methods allow near-optimal matching of most source and channel codes, as well as on-line adaptation to time-varying channels. Techniques that minimize the total system power have also been developed. Developed optimal methods for multi-level broadcasting, joint source-network coding for video transmission over wireless networks and the internet, and total system optimization for distributed wireless sensor networks.

- **Wireless communication for binaural hearing aids.** Developing new methods for low-power wireless communication for binaural hearing aids and other near-the-body applications.

- **Nonstationary interference cancellation.** Developed new frame-based methods for blind removal of nonstationary interference from direct-sequence spread-spectrum communications signals.

- **Stochastic sensor networks.** Developing a very simple, robust, low-power sensor network approach and protocol for large networks of sensors. Each sensor controls power by operating, independently of the others, on a low duty cycle. Recent results prove that such a network can operate reliably with very high probability and can perform all network functions without any coordination of wake/sleep cycle between the nodes.

Low-Power Computer Systems:

- **Global Resource Allocation through Cooperation (GRACE).** In collaboration with Profs. S. Adve, R. Kravets, and K. Nahrstedt in the Computer Science Department, we are developing a new computing framework allowing joint, cooperative adaptation of the hardware, networking, operating system, and media application software to jointly minimize the **total** energy consumption in power-limited, mobile, general-purpose computers.

Signal processing algorithms and implementations:

- **FFTs and Hartley transforms.** Performed the first accurate analysis of the computational complexity of the Hartley transform, showing conclusively that it is virtually equivalent to the real-valued FFT. Co-authored two heavily-cited papers on the Hartley transform and real-valued FFTs.
- **Joint hardware/algorithm design.** Developed several new algorithms/architectures for FFTs, FIR filters, and adaptive filters offering higher performance or reduced hardware complexity.

MAJOR ACCOMPLISHMENTS IN TEACHING

- **Digital Signal Processing Laboratory Textbook.** Developed the first textbook for a DSP-microprocessor-based laboratory course. This text spurred the development of hands-on DSP laboratory courses at many universities, and similar courses are now a mainstay of many electrical and computer engineering curricula in the U.S. and around the world.

  Completed the first open-source, on-line DSP laboratory textbook as part of the Connexions project (`http://cnx.rice.edu`).

- **ECE 320: Digital Signal Processing Laboratory.** Introduced a digital signal processing laboratory at the University of Illinois in 1989. Innovations in both the content and the teaching methods keep this laboratory at the forefront of hands-on DSP education. During the next few years, this laboratory and the students will participate in a NSF-supported research project for the development of next-generation DSP compiler technology. Over the years, equipment donations from Texas Instruments, Motorola, and Analog Devices have equipped the instructional laboratory with state-of-the-art real-time DSP platforms. This course has become one of the most popular laboratory courses in the ECE department (capped at 60 students per semester due to space limitations) and is now offered every semester. ECE 320 has been designated a Texas Instruments Elite Laboratory.

- **ECE 210: Analog Signal Processing.** Assisted in the development of the course outline and the laboratories for this innovative sophomore course, which replaces the traditional sophomore circuits course and the junior signals and systems course. Taught the first full-scale offering of ECE 210 (110+ students). This course is required for all EE and CE majors in the current undergraduate curricula at Illinois.

- **Nonmajors course on Information Technology.** With Prof. Michael Loui, developed a course, to be ECE 101, on information technology and engineering for non-engineering students. Digital information technology is introduced at several leve ls, including audio and video media, digital logic, and the Internet. The course is half lab-based, with students doing real engineering designs so that they also learn the process of technology development and the tradeoffs faced by enginee rs. The course satisfies the General Education requirements for non-science-and-engineering students.

- **Curriculum development.** Served as secretary of the subcommittee which drafted the new Electrical Engineering curriculum at the University of Illinois, and actively involved in current curricular revision.

- **Laboratory enhancement.** Secured a \$71,980 equipment gift from Hewlett-Packard for equipping communications-related laboratories in the department.

PUBLICATIONS

Books

1. Jones, D.L. and T.W. Parks, *A Digital Signal Processing Laboratory for the TMS32010*, Prentice-Hall, Englewood Cliffs, New Jersey, 1988.

2. Appadwedula, S., M.J. Berry, M. Butala, M.A. Haun, D.L. Jones, M.L. Kramer, D. Moussa, D.G. Sachs, B. Wade, *DSP Laboratory with TI TMS320C54x*, Connexions Project, 2002. `http://cnx.rice.edu/content/col10078/latest/`

Book Chapters

1. Jenkins, W.K., B.J. Hunsinger and D.L. Jones, "Discrete-Time Signal Processing," in *Reference Data for Engineers: Radio, Electronics, Computer, and Communications, Eighth Edition*, Edward C. Jordan, Editor, Howard W. Sams & Co., Inc., 1992.

2. Marple, S.L., T. Brotherton and D.L. Jones, "The Application of Advanced Spectral Estimation Techniques to Doppler Ultrasound," in *Time-Frequency and Wavelets in Biomedical Engineering*, Metin Akay, Editor, IEEE Press, 1996.

3. Jenkins, W.K., D.L. Jones, and B.J. Hunsinger, "Discrete-Time Signal Processing," Chapter 25 in *Reference Data for Engineers: Radio, Electronics, Computer, and Communications, Ninth Edition*, Mac E. van Valkenburg and W.M. Middleton, Editors, Howard W. Sams and Co., 2002.

4. Rao, A.M. and D.L. Jones, "Quadratic Detection in Arrays using TFDs," Chapter 8.3 in *Time-Frequency Signal Analysis and Processing: A Comprehensive Reference*, B. Boashash, ed., Elsevier Ltd., Oxford, UK, 2003, pp. 344–347. (ISBN: 0-08-044335-4)

5. Baraniuk, R.G. and D.L. Jones, "Adaptive Time-Frequency Representations," Chapter 5.3 in *Time-Frequency Signal Analysis and Processing: A Comprehensive Reference*, B. Boashash, ed., Elsevier Ltd., Oxford, UK, 2003, pp. 178–184. (ISBN: 0-08-044335-4)

6. Kwok, H.K. and D.L. Jones, "Shell Mapping," in *Encyclopedia of Telecommunications*, J.G. Proakis, ed., John Wiley, 2003.

Journal Publications

1. Sorensen, H.V., D.L. Jones, C.S. Burrus and M.T. Heideman, "On Computing the Discrete Hartley Transform," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. 33, no. 5, October 1985, pp. 849–863.

2. Sorensen, H.V., D.L. Jones, M.T. Heideman and C.S. Burrus, "Real-Valued Fast Fourier Transform Algorithms," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, , vol. 35, no. 6, June 1987, pp. 1231–1238, and "Corrections to 'Real-Valued Fast Fourier Transform Algorithms'," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. 35, no. 9, September 1987, p. 1353.

3. Jones, D.L. and T.W. Parks, "Time-Frequency Window Leakage in the Short–Time Fourier Transform," *Circuits, Signals, and Signal Processing*, vol. 6, no. 3, 1987, pp. 263–286.

4. Jones, D.L. and T.W. Parks, "On Computing Equally Spaced Samples of a Complex Gaussian Function," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. 35, no. 10, October 1987, pp. 1479–1480.

5. Jones, D.L., G.L. Hoehn and A.F. Kuckes, "Improved Magnetic Model for Determination of Range and Direction to a Blowout Well," *SPE Drilling Engineering*, December 1987, pp. 316–322.

6. Jones, D.L. and T.W. Parks, "On the Generation and Combination of Grains for Music Synthesis," *Computer Music Journal*, vol. 12, no. 2, Summer 1988, pp. 27–34.

7. Jones, D.L. and T.W. Parks, "A High Resolution Data-Adaptive Time-Frequency Representation," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. 38, no. 12, December 1990, pp. 2127–2135.

8. Jones, D.L. and R.G. Baraniuk, "Efficient Approximation of Continuous Wavelet Transforms," *Electronics Letters*, vol. 27, no. 9, April 25, 1991, pp. 748–750.

9. Jones, D.L. and H.V. Sorensen, "A Bus-Oriented Multiprocessor Fast Fourier Transform," *IEEE Trans. on Signal Processing*, vol. 39, no. 11, November 1991, pp. 2547–2552.

10. Jones, D.L. and T.W. Parks, "A Resolution Comparison of Several Time-Frequency Representations," *IEEE Trans. on Signal Processing*, vol. 40, no. 2, February 1992, pp. 413–420.

11. Jones, D.L., "Learning Characteristics of Transpose-Form LMS Adaptive Filters," *IEEE Trans. on Circuits and Systems Part II: Analog and Digital Signal Processing*, vol. 39, no. 10, October 1992, pp. 745–749.

12. Jones, D.L., "Efficient Computation of Time-Varying and Adaptive Filters," *IEEE Trans. on Signal Processing*, vol. 41, no. 3, March 1993, pp. 1077–1086.

13. Baraniuk, R.G. and D.L. Jones, "A Signal-Dependent Time-Frequency Representation: Optimal Kernel Design," *IEEE Trans. on Signal Processing*, vol. 41, no. 4, April 1993, pp. 1589–1602.

14. Baraniuk, R.G. and D.L. Jones, "Signal-Dependent Time-Frequency Analysis Using a Radially-Gaussian Kernel," *Signal Processing*, vol. 32, no. 3, May 1993, pp. 263–284.

15. Jones, D.L., J.S. Touvannas, P. Lander and D.E. Albert, "Advanced Time-Frequency Methods for Signal-Averaged Electrocardiogram Analysis," *Journal of Electrocardiology*, vol. 25, Supplement, 1993, pp. 188–194.

16. Miyazu, K. and D.L. Jones, "An Adaptive FIR Filtering Using Acoustic Charge Transport Device with Bypass Digital Filter," *IEICE Transactions on Fundamentals of Electronics, Communications, and Computer Sciences*, vol. E76-A, no. 8, August 1993, pp. 1323–1325.

17. Baraniuk, R.G. and D.L. Jones, "Shear Madness: New Orthonormal Bases and Frames Using Chirp Functions," *IEEE Trans. on Signal Processing*, vol. 41, no. 12, December 1993, pp. 3543–3549.

18. Baraniuk, R.G. and D.L. Jones, "A Signal-Dependent Time-Frequency Representation: Fast Algorithm for Optimal Kernel Design," *IEEE Trans. on Signal Processing*, vol. 42, no. 1, January 1994, pp. 134–146.

19. Jones, D.L. and R.G. Baraniuk, "A Simple Scheme for Adapting Time-Frequency Representations," *IEEE Trans. on Signal Processing*, vol. 42, no. 12, December 1994, pp. 3530–3535.

20. Sayeed, A.M. and D.L. Jones, "Optimal Kernels for Non-Stationary Spectral Estimation," *IEEE Trans. on Signal Processing*, vol. 43, no. 2, February 1995, pp. 478–491.

21. Czerwinski, R.N. and D.L. Jones, "Adaptive Cone Kernel Time-Frequency Analysis," *IEEE Trans. on Signal Processing*, vol. 43, no. 7, July 1995, pp. 1715–1719.

22. Baraniuk, R.G. and D.L. Jones, "Unitary Equivalence: A New Twist on Signal Processing," *IEEE Trans. on Signal Processing*, vol. 43, no. 10, October 1995, pp. 2269–2282.

23. Jones, D.L. and R.G. Baraniuk, "An Adaptive Optimal-Kernel Time-Frequency Representation," *IEEE Trans. on Signal Processing*, vol. 43, no. 10, October 1995, pp. 2361–2371.

24. Sayeed, A.M. and D.L. Jones, "Optimal Detection Using Bilinear Time-Frequency and Time-Scale Representations," *IEEE Trans. on Signal Processing*, vol. 43, no. 12, December 1995, pp. 2872–2883.

25. Sayeed, A.M., P. Lander and D.L. Jones, "Improved Time-Frequency Filtering of Signal-Averaged ECGs," *Journal of Electrocardiology*, vol. 28, Supplement, 1995, pp. 53–58.

26. Sayeed, A.M. and D.L. Jones, "A Canonical Covariance-Based Method for Generalized Joint Signal Representations," *IEEE Signal Processing Letters*, vol. 3, no. 4, April 1996, pp. 121–123.

27. Sayeed, A.M. and D.L. Jones, "Integral Transforms Covariant to Unitary Operators and their Implications for Joint Signal Representations," *IEEE Trans. on Signal Processing*, vol. 44, no. 6, June 1996, pp. 1365–1377.

28. Sayeed, A.M. and D.L. Jones, "Equivalence of Generalized Joint Signal Representations of Arbitrary Variables," *IEEE Trans. on Signal Processing*, vol. 44, no. 12, December 1996, pp. 2959–2970.

29. Sayeed, A.M. and D.L. Jones, "Optimal Quadratic Detection and Estimation Using Generalized Joint Signal Representations," *IEEE Trans. on Signal Processing*, vol. 44, no. 12, December 1996, pp. 3031–3043.

30. Baraniuk, R.G. and D.L. Jones, "Wigner-Based Formulation of the Chirplet Transform," *IEEE Trans. on Signal Processing*, vol. 44, no. 12, December 1996, pp. 3129–3135.

31. Czerwinski, R.N. and D.L. Jones, "Adaptive Short-Time Fourier Analysis," *IEEE Signal Processing Letters*, vol. 4, no. 2, February 1997, pp. 42–45.

32. Gomis, P., D.L. Jones, P. Caminal, E.J. Berbari, and P. Lander, "Analysis of Abnormal Signals Within the QRS Complex of the High-Resolution Electrocardiogram," *IEEE Transactions on Biomedical Engineering*, vol. 44, no. 8, August 1997, pp. 681–693.

33. Yang, V.Y. and D.L. Jones, "A Vector Constant Modulus Algorithm for Shaped Constellation Equalization," *IEEE Signal Processing Letters*, vol. 5, no. 4, April 1998, pp. 89–91.

34. Czerwinski, R.N., D.L. Jones and W.D. O'Brien, Jr., "Line and Boundary Detection in Speckle Images," *IEEE Trans. on Image Processing*, vol. 7, no. 12, December 1998, pp. 1700–1714.

35. Czerwinski, R.N., D.L. Jones and W.D. O'Brien, Jr., "Detection of Lines and Curves in Speckle Images — Application to Ultrasound Imaging," *IEEE Trans. on Medical Imaging*, vol. 18, no. 2, February 1999, pp. 126–136.

36. Krongold, B.S., A.M. Sayeed, M. Moehring, J.A. Ritcey, M. Spencer and D.L. Jones, "Time-Scale Detection of Microemboli in Flowing Blood with Doppler Ultrasound," *IEEE Transactions on Biomedical Engineering*, vol. 46, no. 9, September 1999, pp. 1081–1089.

37. Krongold, B.S., K. Ramchandran and D.L. Jones, "Frequency-Shift-Invariant Orthonormal Wavelet Packet Representations," *IEEE Transactions on Signal Processing*, vol. 47, no. 9, September 1999, pp. 2579–2582.

38. Munson, D.C. Jr. and D.L. Jones, "Analog Signal Processing First," *IEEE Signal Processing Magazine*, vol. 16, no. 5, September 1999, pp. 24–29.

39. Krongold, B.S., K. Ramchandran and D.L. Jones, "Computationally Efficient Optimal Power Allocation Algorithms for Multicarrier Communication Systems," *IEEE Transactions on Communications*, vol. 48, no. 1, January 2000, pp. 23–27.

40. Rao, A.M. and D.L. Jones, "A Denoising Approach to Multichannel Signal Estimation," *IEEE Transactions on Signal Processing*, vol. 48, no. 5, May 2000, pp. 1225–1234.

41. Liu, C., B.C. Wheeler, W.D. O'Brien, Jr., R.C. Bilger, D.L. Jones, C.R. Lansing, A.S. Feng, "Localization of Multiple Sound Sources with Two Microphones in Complex Acoustic Environments," *Journal of the Acoustical Society of America*, vol. 108, no. 4, 2000, pp. 1888–1905.

42. Schwartz, B.C. and D.L. Jones, "Quadratic and Instantaneous Frequency Analysis of Helicopter Gearbox Faults," *Mechanical Systems and Signal Processing*, vol. 14, no. 4, July 2000, pp. 579–595.

43. Kwok, H.K-C. and D.L. Jones, "Improved Instantaneous Frequency Estimation Using an Adaptive Short-Time Fourier Transform," *IEEE Trans. on Signal Processing*, vol. 48, no. 10, October 2000, pp. 2964–2972.

44. Appadwedula, S., M. Goel, N.R. Shanbhag, D.L. Jones, and K. Ramchandran, "Total System Energy Minimization for Wireless Image Transmission," *Journal of VLSI Signal Processing Systems*, vol. 27, no. 1, February 2001, pp. 99–117.

45. Liu, C., B.C. Wheeler, W.D. O'Brien, Jr., C.R. Lansing, R.C. Bilger, D.L. Jones, A.S. Feng, "A two-microphone dual delay-line approach for extraction of a speech sound in the presence of multiple interferers," *Journal of the Acoustical Society of America*, vol. 110, no. 6, 2001, pp. 3218–3231.

46. Rao, A.M. and D.L. Jones, "Efficient Quadratic Detection in Perturbed Arrays via Fourier Transform Techniques," *IEEE Transactions on Signal Processing*, vol. 49, no. 7, July 2001, pp. 1269–1281.

47. Haun, M.A., D.L. Jones, and W.D. O'Brien, Jr., "Efficient Three-Dimensional Imaging from a Small Cylindrical Aperture," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 49, no. 7, July 2002, pp. 861–870.

48. Rao, A.M. and D.L. Jones, "Efficient Detection with Arrays in the Presence of Angular Spreading," *IEEE Transactions on Signal Processing*, vol. 51, no. 2, February 2003, pp. 301–312.

49. Krongold, B.S. and D.L. Jones, "Peak Power Reduction via Active Constellation Extension," *IEEE Transactions on Broadcasting*, vol. 49, no. 3, September 2003, pp. 258–268.

50. Ratnam, R, D.L. Jones, B.C. Wheeler, W.D. O'Brien, Jr., C.R. Lansing, and A.S. Feng, "Blind Estimation of Reverberation Time," *Journal of the Acoustical Society of America*, vol. 114, no. 5, November 2003, pp. 2877-2892.

51. Lockwood, M.E., Jones, D.L., Bilger, R.C., Lansing, C.R., O'Brien, W.D., Wheeler, B.C. and Feng, A.S., "Performance of time- and frequency-domain binaural beamformers based on recorded signals from real rooms," *Journal of the Acoustical Society of America*, vol. 115, no. 1, January 2004, pp. 379–391.

52. Krongold, B.S. and D.L. Jones, "An Active-Set Approach for OFDM PAR Reduction via Tone Reservation," *IEEE Transactions on Signal Processing*, vol 52, no. 2, February 2004, pp. 495–509.

53. Ratnam, R. and D.L. Jones, and W.D. O'Brien, "Fast algorithm for blind reverb time estimation," *IEEE Signal Processing Letters*, vol. 11, no. 6, June 2004, pp. 537–540.

54. Baraniuk, R.G., C.S. Burrus, D.H. Johnson, and D.L. Jones, "Sharing Knowledge and Building Communities in Signal Processing," *IEEE Signal Processing Magazine*, vol. 21, no. 5, September 2004, pp. 10–16.

55. Haun, M.A., D.L. Jones, and W.D. O'Brien, Jr., "Overdetermined Least-Squares Aberration Estimates Using Common-Midpoint Signals," *IEEE Transactions on Medical Imaging*, vol. 23, no. 10, October 2004, pp. 1205–1220.

56. Appadwedula, S., V.V. Veervalli, and D.L. Jones, "Energy-Efficient Detection in Sensor Networks," to appear in the *IEEE Journal on Selected Areas in Communications*, 2005.

Conference Proceedings

1. Jones, D.L., G.L. Hoehn and A.F. Kuckes, "Improved Magnetic Model for Determination of Range and Direction to a Blowout Well," SPE 14388, *Soc. of Petr. Engineers Annual Conference*, Las Vegas, Nevada, September 22–25, 1985.

2. Sorensen, H.V., D.L. Jones, M.T. Heideman and C.S. Burrus, "A Split-Radix Real-Valued Fourier Transform," *Proc. European Signal Processing Conference EUSIPCO-86*, The Hague, The Netherlands, September 2–5, 1986.

3. Jones, D.L. and T.W. Parks, "Time-Scale Modification Using a Synchronous Gabor Technique," *Proc. IEEE 1986 Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, New York, September 14–17, 1986.

4. Jones, D.L. and T.W. Parks, "Matched Windows in the Short-Time Fourier Transform," *Proc. IEEE ASSP 1986 Digital Signal Processing Workshop*, Chatham, Massachusetts, October 20–22, 1986, pp. 2.2.1-2.2.2.

5. Jones, D.L. and T.W. Parks, "A High Resolution Data-Adaptive Time-Frequency Representation," *Proc. 1987 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Dallas, Texas, April 6–9, 1987, pp. 681–684.

6. Sorensen, H.V., D.L. Jones and C.S. Burrus, "Real-Valued Algorithms for the FFT," *Proc. 1987 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Dallas, Texas, April 6–9, 1987, pp. 1831–1834.

7. Sorensen, H.V., C.S. Burrus and D.L. Jones, "A New Efficient Algorithm for Computing a Few DFT Points," *Proc. 1988 IEEE International Symposium on Circuits and Systems*, Helsinki, Finland, June 7–9, 1988, pp. 1915–1918.

8. Jones, D.L. and H.V. Sorensen, "A Bus-Oriented Multiprocessor Fast Fourier Transform," *Proc. 1989 Conf. on Information Science and Systems*, Baltimore, Maryland, March 22–24 1989, pp. 631–636.

9. Jones, D.L. and T.W. Parks, "A Resolution Comparison of Several Time-Frequency Representations,"*Proc. IEEE 1989 International Conference on Acoustics, Speech, and Signal Processing*, Glasgow, Scotland, U.K., May 23–26, 1989, pp. 2222–2225.

10. Young, C. and D.L. Jones, "Area-Efficient VLSI Implementation of Digital Filters via Multiple Product Intercoding," *Proc. 32nd Midwest Symposium on Circuits and Systems*, Champaign, Illinois, August 14–16, 1989, pp. 373–376.

11. Jones, D.L, "Efficient Scalar-Vector Product Computation for Signal Processing," *Proc. 23rd Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Oct. 30–Nov. 1, 1989, pp. 28–32.

12. Koczwara, T.E. and D.L. Jones, "On Mask Selection for Time-Varying Filtering Using the Wigner Distribution," *Proc. IEEE 1990 International Conference on Acoustics, Speech, and Signal Processing*, Albuquerque, New Mexico, April 3–6, 1990, pp. 2487–2490.

13. Baraniuk, R.G. and D.L. Jones, "Optimal Kernels for Time-Frequency Analysis," *Advanced Signal-Processing Algorithms, Architectures, and Implementations*, Franklin T. Luk, Editor, Proc. SPIE 1348, San Diego, California, July 10–12, 1990, pp. 181–187.

14. Jones, D.L. and R.G. Baraniuk, "Signal Dependent Time-Frequency Representations," *Proc. IEEE Fourth Digital Signal Processing Workshop*, New Paltz, New York, September 16–19, 1990, pp. 3.15.1-3.15.2.

15. Potter, L.C., K.S. Arun and D.L. Jones, "Recovery of Pore-Size Distributions from NMR Spin-Lattice Relaxation Measurements," *Proc. IEEE Fourth Digital Signal Processing Workshop*, New Paltz, New York, September 16–19, 1990, pp. 2.1.1-2.1.2.

16. Baraniuk, R.G. and D.L. Jones, "A Radially Gaussian Signal-Dependent Kernel for Time-Frequency Analysis," *Proc. IEEE 1991 International Conference on Acoustics, Speech, and Signal Processing*, Toronto, Canada, May 14–17, 1991, pp. 3181–3184.

17. Jones, D.L. and R.G. Baraniuk, "Efficient Computation of Densely Sampled Wavelet Transforms," *Advanced Signal-Processing Algorithms, Architectures, and Implementations II*, Franklin T. Luk, Editor, Proc. SPIE 1566, San Diego, California, July 21–26, 1991, pp. 202-206.

18. Young, C. and D.L. Jones, "Area-Efficient VLSI Implementation of FIR Digital Filters Using Signed Digit Coding," *Proc. 25th Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 4-6, 1991, pp. 398–402.

19. Baraniuk, R.G., D.L. Jones, T. Brotherton and S.L. Marple, "Applications of Adaptive Time-Frequency Representations to Underwater Acoustic Signal Processing," *Proc. 25th Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 4-6, 1991, pp. 1109–1113.

20. Baraniuk, R.G. and D.L. Jones. "New Dimensions in Wavelet Analysis," *Proc. IEEE 1992 International Conference on Acoustics, Speech, and Signal Processing*, San Francisco, March 23–26, 1992, pp. V-137–V-140.

21. Young, C. and D.L. Jones. "Improvement in Finite Wordlength FIR Digital Filter Design," *Proc. IEEE 1992 International Conference on Acoustics, Speech, and Signal Processing*, San Francisco, March 23–26, 1992, pp. V-109–V-112.

22. Ezell, J.E., W.K. Jenkins, D.L. Jones, K. Miyazu, G.J. Siampos and G.E. Stanton, "Adaptive Analog Signal Processing with Acoustic Charge Transport Filters," *Proc. IEEE 1992 International Symposium on Circuits and Systems*, San Diego, May 10–13, 1992.

23. Jones, D.L., "Advanced Spectral Estimation Methods," *17th Annual Conference on Research and Applications in Computerized Electrocardiology*, Keystone, Colorado, May 2–7, 1992.

24. Jones, D.L. and R.G. Baraniuk, "An On-Line Signal-Dependent Time-Frequency Representation," *Proc. IEEE Fifth Digital Signal Processing Workshop*, Starved Rock State Park, Illinois, September 13–16, 1992, pp. 1.3.1-1.3.2.

25. Baraniuk, R.G. and D.L. Jones, "New Signal-Space Orthogonal Bases via the Metaplectic Transform," *Proc. IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, Victoria, British Columbia, October 4–6, 1992, pp. 339–342.

26. Jones, D.L. and R.G. Baraniuk, "A Simple Scheme for Adapting Time-Frequency Representations," *Proc. IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, Victoria, British Columbia, October 4–6, 1992, pp. 83–86.

27. Brotherton, T., T. Pollard and D.L. Jones, "Applications of Time-Frequency and Time-Scale Representations to Fault Detection and Classification," *Proc. IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, Victoria, British Columbia, October 4–6, 1992, pp. 95–98.

28. Jones, D.L. and R.G. Baraniuk, "An Adaptive Time-Frequency Representation with a Radially Gaussian Kernel," *Proc. IEEE 1993 International Conference on Acoustics, Speech, and Signal Processing*, Minneapolis, Minnesota, April 27–30, 1993, pp. IV-109–112.

29. Baraniuk, R.G. and D.L. Jones, "Warped Wavelet Bases: Unitary Equivalence and Signal Processing," *Proc. IEEE 1993 International Conference on Acoustics, Speech, and Signal Processing*, Minneapolis, Minnesota, April 27–30, 1993, pp. III-320–323.

30. Czerwinski, R.N. and D.L. Jones, "An Adaptive Time-Frequency Representation Using a Cone-Shaped Kernel," *Proc. IEEE 1993 International Conference on Acoustics, Speech, and Signal Processing*, Minneapolis, Minnesota, April 27–30, 1993, pp. IV-104–107.

31. Canfield, K.G. and D.L. Jones, "Implementing Time-Frequency Representations for non-Cohen Classes," *Proc. 27th Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 1-3, 1993, pp. 1464–1468.

32. Marple, L.S., T. Brotherton, R. Barton, K. Lugo, and D.L. Jones, "Travels Through the Time-Frequency Zone: Advanced Doppler Ultrasound Processing Techniques," *Proc. 27th Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 1-3, 1993, pp. 1469–1474.

33. Czerwinski, R.N., D.L. Jones, and W.D. O'Brien, Jr., "An Approach to Boundary Detection in Ultrasound Imaging," *Proc. 1993 IEEE Ultrasonics Symposium*, Baltimore, MD, Oct. 31-Nov. 3, 1993, pp. 951–955.

34. Sayeed, A.M. and D.L. Jones, "Optimal Kernels for Wigner-Ville Spectral Estimation," *Proc. IEEE 1994 International Conference on Acoustics, Speech, and Signal Processing*, Adelaide, Australia, April 19–22, 1994, pp. IV-297–300.

35. Kramer, M.L. and D.L. Jones, "Improved Time-Frequency Filtering Using an STFT Analysis-Modification-Synthesis Method," *Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, Philadelphia, October 25–28, 1994, pp. 264–267.

36. Sayeed, A.M. and D.L. Jones, "Optimal Quadratic Detection Using Bilinear Time-Frequency and Time-Scale Representations," *Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, Philadelphia, October 25–28, 1994, pp. 365–368.

37. Sayeed, A.M. and D.L. Jones, "Optimal Estimation of Time-Frequency Representations from Corrupted Observations," *Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, Philadelphia, October 25–28, 1994, pp. 456–459.

38. Czerwinski, R.N., D.L. Jones, and W.D. O'Brien, Jr., "Edge Detection in Ultrasound Speckle Noise," *Proceedings of the IEEE International Conference on Image Processing*, vol. III, pp. 304–308, Austin, Texas, November 13-16, 1994.

39. Sayeed, A.M. and D.L. Jones, "Improved Time-Frequency Filtering of Signal-Averaged Electrocardiograms," presented at the *20th Annual Conference on Research and Applications in Computerized Electrocardiography*, Amelia Island, Florida, May 1995.

40. Jones, D.L. and A.M. Sayeed, "Blind Quadratic and Time-Frequency Based Detectors from Training Data," *Proc. IEEE Int. Conf. on Acoust., Speech and Signal Proc. — ICASSP '95*, Detroit, Michigan, May 8-12, 1995, pp. 1033–1036.

41. Sayeed, A.M. and D.L. Jones, "On the Equivalence of Generalized Joint Signal Representations," *Proc. IEEE Int. Conf. on Acoust., Speech and Signal Proc. — ICASSP '95*, Detroit, Michigan, May 8-12, 1995, pp. 1533–1536.

42. Czerwinski, R.N., D.L. Jones, and W.D. O'Brien, Jr., "A Comparison of Detection Rules for Lines and Boundaries in Acoustic Speckle," *22nd International Symposium on Acoustical Imaging*, Florence, Italy, September 1995.

43. Czerwinski, R.N., D.L. Jones, and W.D. O'Brien, Jr., "Ultrasound Speckle Reduction by Directional Median Filtering," *Proceedings of the IEEE International Conference on Image Processing*, vol. I, pp. 358–361, Washington, DC, October 23-26, 1995.

44. Kwok, H. K.-C. and D.L. Jones, "Instantaneous Frequency Estimation Using an Adaptive Short-Time Fourier Transform," *Proc. 29th Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Oct. 29–Nov. 1, 1995, pp. 543–547.

45. Jones, D.L., "A Normalized Constant Modulus Algorithm," *Proc. 29th Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Oct. 29–Nov. 1, 1995, pp. 694–697.

46. Czerwinski, R.N., D.L. Jones, and W.D. O'Brien, Jr., "Decision Directed Line Detection with Application to Medical Ultrasound," *1996 SPIE International Symposium on Medical Imaging*, Newport Beach, CA., February 1996.

47. Sayeed, A.M. and D.L. Jones, "Time-Frequency Detectors," *Proc. 1996 Conf. on Information Sciences and Systems*, Princeton, NJ, vol. 1, March 1996, pp. 509–514.

48. Kramer, M.L. and D.L. Jones, "A Pitch-Dependent Time-Frequency Response for Periodic Signals," *Proc. IEEE Int. Conf. on Acoust., Speech and Signal Proc. — ICASSP '96*, Atlanta, Georgia, May 7-10, 1996, pp. IV-2598–2601.

49. Sayeed, A.M. and D.L. Jones, "Generalized Joint Signal Representations and Optimum Detection," *Proc. IEEE Int. Conf. on Acoust., Speech and Signal Proc. — ICASSP '96*, Atlanta, Georgia, May 7-10, 1996, pp. III-1431-1434.

50. Samimy, B., G. Rizzoni, A.M. Sayeed and D.L. Jones, "Design of Training Data-Based Quadratic Detectors with Application to Mechanical Systems," *Proc. IEEE Int. Conf. on Acoust., Speech and Signal Proc. — ICASSP '96*, Atlanta, Georgia, May 7-10, 1996, pp. III-1767–1770.

51. Kwok, H. K.-C. and D.L. Jones, "Improved FM Demodulation in a Fading Environment," *Proc. IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, Paris, France, June 18-21, 1996, pp. 9–12.

52. Sayeed, A.M. and D.L. Jones, "Optimal Reduced-Rank Time-Frequency/Time-Scale Quadratic Detectors," *Proc. IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, Paris, France, June 18-21, 1996, pp. 209–212.

53. Sayeed, A.M. and D.L. Jones, "A Simple Covariance-Based Characterization of Joint Signal Representations of Arbitrary Variables," *Proc. IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, Paris, France, June 18-21, 1996, pp. 433–436.

54. Czerwinski, R.N., D.L. Jones, and W.D. O'Brien, Jr., "Linear Feature Enhancement by False Color," *Proceedings of the 1996 IEEE Ultrasonics Symposium*, San Antonio, TX, November 1996.

55. Krongold, B.S., M.L. Kramer, K. Ramchandran, and D.L. Jones, "Spread Spectrum Interference Suppression Using Adaptive Time-Frequency Tilings," *Proc. IEEE Int. Conf. on Acoust., Speech and Signal Proc. — ICASSP '97*, Munich, Germany, April 21-24, 1997, pp. 1881–1884.

56. Krongold, B.S., A.M. Sayeed, M. Moehring, J.A. Ritcey, M. Spencer, and D.L. Jones, "Quadratic Time-Scale Detection of Microemboli in Flowing Blood With Doppler Ultrasound," *Advanced Signal-Processing Algorithms, Architectures, and Implementations VII*, Franklin T. Luk, Editor, Proc. SPIE 3162, San Diego, California, July 27–August 1, 1997, pp. 137–148.

57. Nollett, B.S., and D.L. Jones, "Nonlinear Echo Cancellation for Hands-Free Speakerphones," *Proc. 1997 IEEE Workshop on Nonlinear Signal and Image Processing*, Mackinac Island, Michigan, Sept. 8–10, 1997.

58. Rao, A.M., and D.L. Jones, "Nonstationary Array Signal Detection and Source Localization Using Time-Frequency and Time-Scale Representations," *5th Battlefield Acoustics Symposium*, Ft. Meade, Maryland, Sept. 23–25, 1997.

59. Krongold, B.S., K. Ramchandran, and D. L. Jones, "Frequency-Shift-Invariant Orthonormal Wavelet Packet Representations," *Proc. IEEE International Conference on Image Processing*, Santa Barbara, CA, 1997, pp. 628–631.

60. Kramer, M.L., and D.L. Jones, "Nonstationary Interference Suppression Using Adaptive Overdetermined Frame Representations," *Proc. 31st Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 2-Nov. 5, 1997.

61. Yang, V.Y., and D.L. Jones, "A Vector Constant Modulus Algorithm for Shaped Constellation Equalization," *Proc. 31st Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 2–Nov. 5, 1997.

62. Appadwedula, S., D.L. Jones, K. Ramchandran, and I. Kozintsev, "Joint Source Channel Matching for a Wireless Communications Link," *Proceedings of the Data Compression Conference*, Snowbird, Utah, March 29–April 1, 1998, p. 523.

63. Rao, Anil M., and D.L. Jones, "Nonstationary Array Signal Detection Using Time-Frequency and Time-Scale Representations," *Proc. IEEE Int. Conf. on Acoust., Speech and Signal Proc. — ICASSP '98*, Seattle, Washington, May 12–15, 1998, pp. 1989–1992.

64. Munson, D.C., and D.L. Jones, "Analog Signal Processing: A Better Way to Teach Circuits and Systems," *IEEE International Symposium on Circuits and Systems*, Monterey, California, May 1998.

65. Krongold, B.S., K. Ramchandran, and D. L. Jones, "Computationally Efficient Optimal Power Allocation Algorithm for Multicarrier Communication Systems," *Proc. International Conference on Communications*, Atlanta, GA, June 7–11, 1998, pp. 1018–1022.

66. Appadwedula, S., D.L. Jones, K. Ramchandran, and I. Kozintsev, "Joint Source-Channel Matching for a Wireless Communications Link," *Proc. International Conference on Communications*, Atlanta, GA, June 7–11, 1998.

67. Appadwedula, S., D.L. Jones, K. Ramchandran, and L. Qian, "Joint Source-Channel Matching for a Wireless Image Transmission," *Proc. International Conference on Image Processing*, Chicago, IL, Oct. 4–7, 1998.

68. Rao, A.M. and D.L. Jones, "Efficient Structures for Quadratic Time-Frequency and Time-Scale Array Processors," *Proc. Intl. Symp. on Time-Frequency and Time-Scale Analysis*, Pittsburgh, PA, Oct. 6–9, 1998, pp. 397–400.

69. Appadwedula, S., M. Goel, D. L. Jones, K. Ramchandran and N. R. Shanbhag, "Efficient Wireless Image Transmission Under a Total Power Constraint," *IEEE Multimedia Signal Processing Workshop*, Redondo Beach, CA, December 1998, pp. 573–578.

70. Qian, L., D.L. Jones, K. Ramchandran, and S. Appadwedula, "Joint Source-Channel Matching for Wireless Video Transmission," *Proc. 5th Intl. Symposium on DSP for Communication Systems — DSPCS'99*, Perth, Australia, February 1–5, 1999, pp. 274–280.

71. Jones, D.L., "A New Method for Blind Source Separation of Nonstationary Signals," *IEEE International Conf. on Acoustics, Speech, and Signal Processing*, Phoenix, AZ, March 15-19, 1999.

72. Munson, D.C. Jr., and D.L. Jones, "Analog Signal Processing First," *IEEE International Conf. on Acoustics, Speech, and Signal Processing*, Phoenix, AZ, March 15-19, 1999.

73. Rao, A.M., and D.L. Jones, "A Denoising Approach to Multichannel Signal Estimation," *IEEE International Conf. on Acoustics, Speech, and Signal Processing*, Phoenix, AZ, March 15-19, 1999, pp. 2869–2872.

74. Haun, M., and D.L. Jones, "The Fractionally Spaced Vector Constant Modulus Algorithm," *IEEE International Conf. on Acoustics, Speech, and Signal Processing*, Phoenix, AZ, March 15-19, 1999.

75. Qian, L., D.L. Jones, and K. Ramchandran, and S. Appadwedula, "A General Joint Source-Channel Matching Method for Wireless Video Transmission," *Proceedings of the Data Compression Conference*, Snowbird, Utah, March 29–April 1, 1999, pp. 414–423.

76. Jones, D.L., "Effective DSP Laboratory Course Design," *DSPSFest'99*, Houston, Texas, August 4–6, 1999.

77. Jones, D.L., "An Extended Gatherer-Polley Method for OFDM and DMT Peak Power Reduction," *DSPSFest'99*, Houston, Texas, August 4–6, 1999.

78. Jones, D.L., "Peak Power Reduction in OFDM and DMT via Active Channel Modification," *Proc. 33st Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Oct. 24-27, 1999, pp. 1076–1079.

79. Krongold, B.S., K. Ramchandran and D. L. Jones, "An Efficient Algorithm for Optimal Margin Maximization in Multicarrier Communications Systems," *Proceedings of IEEE Globecom '99*, Rio de Janeiro, Brazil, 1999, pp. 899–903.

80. Lockwood, M.E., Jones, D.L., Elledge, M.E., Bilger, R.C., Goueygou, M., Lansing, C.R., Liu, C., O'Brien, W.D. Jr., Wheeler, B.C., Feng, A.S. (1999) A minimum variance frequency-domain algorithm for binaural hearing aid processing. J. Acous. Soc. Am. 106 (Pt. 2): 2278A.

81. Elledge, M.E., Lockwood, M.E., Jones, D.L., Bilger, R.C., Goueygou, M., Lansing, C.R., Liu, C., O'Brien, W.D. Jr., Wheeler, B.C., Feng, A.S. (1999) A real-time dual-microphone signal-processing system for hearing aids. J. Acous. Soc. Am. 106 (Pt. 2): 2279A.

82. D.L. Jones, "Characteristics of Signal Processing Applications and the Potential of Processor-in-Memory Architectures," *International Workshop on Innovative Architecture for Future Generation High-Performance Processors and Systems (IWIA'99)*, Maui, Hawaii, Nov. 1–3, 1999.

83. Kwok, H. K.-C. and D.L. Jones, "PAR Reduction for Hadamard Transform-Based OFDM," *Proc. 2000 Conference on Information Sciences and Systems*, Princeton, New Jersey, March 15-17, 2000, pp. FP4-18–23.

84. Rao, A.M and D.L. Jones, "Denoising Multisensor Data," *Proc. First IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM 2000)*, Cambridge, MA, March 2000, pp. 375–379.

85. Kwok, H. K.-C. and D.L. Jones, "PAR Reduction via Constellation Shaping," *Proc. IEEE International Symposium on Information Theory*, Sorrento, Italy, June 25–30, 2000, p. 166.

86. Jones, D.L., D.S. Roberts and B.S. Krongold, "Biomedical Applications of Time-Frequency and Time-Scale Detection," *IEEE EMBS Workshop on Time-Freq and Time-Scale in Biomedicine*, Chicago, IL, July 22, 2000.

87. Kramer, M.L., M.A. Haun, S. Appadwedula, D.G. Sachs, and D.L. Jones, "Effective Use of Projects in DSP Laboratory Instruction," *DSPSFest'2000*, Houston, Texas, August 2–4, 2000.

88. Elledge, M.E., Lockwood, M.E., Bilger, R.C., Feng, A.S., Jones, D.L., Lansing, C.R. O'Brien, W.D. and Wheeler, B.C., "Real-time implementation of a frequency-domain beamformer on the TI C62X EVM," *DSPSFest'2000*, Houston, Texas, August 2–4, 2000.

89. Kwok, H. K.-C. and D.L. Jones, "An Efficient Power-Reduction Technique for DSL Modems using Constellation Shaping," *DSPSFest'2000*, Houston, Texas, August 2–4, 2000.

90. Rao, A.M and D.L. Jones, "Efficient Signal Detection in Perturbed Arrays," *Proceedings of the IEEE Workshop on Statistical Signal and Array Processing*, Pennsylvania, August 14–16, 2000, pp. 429–433.

91. Krongold, B.S. and D.L. Jones, "Blind Source Separation of Nonstationary Convolutively Mixed Signals," *Proceedings of the IEEE Workshop on Statistical Signal and Array Processing*, Pocono Manor, Pennsylvania, August 14–16, 2000, pp. 53–57.

104. Larsen, J., Lockwood, M.E., Lansing, C.R., Bilger, R.C., O'Brien, W.D., Jones, D.L., Wheeler, B.C. and Feng, A.S. (2001). Performance of a frequency-based minimum variance beamforming algorithm for normal and hearing impaired listeners. J. Acoust. Soc. Am. 109.

105. Kwok, H.C. and D.L. Jones, "Shell Mapping Based Power Reduction Code," *Proc. International Symposium on Information Theory*, Washington, D.C., June 2001.

106. Qian, L. and D.L. Jones, "Minimax Disappointment Criterion for Video Broadcasting," *Proc. of the IEEE International Conference on Image Processing*, Thessaloniki, Greece, October 7–10, 2001.

107. Krongold, B.S. and D.L. Jones, "A New Method for PAR Reduction in Baseband DMT Systems," *Proc. 35th Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, November, 2001, pp. 502–506.

108. Zhao, W.-K., B.-S. Cai, D. Whalley, M. Bailey, R. van Engelen, X. Yuan, J. Hiser, J. Davidson, K. Gallivan, and D.L. Jones, "VISTA: A System for Interactive Code Improvement," Proceedings of the LCTES conference, 2002.

109. Krongold, B.S. and D.L. Jones, "A New Tone Reservation Method for Complex Baseband PAR Reduction in OFDM Systems," ICASSP 2002, Orlando, Florida, May 13-16, 2002.

110. Rao, A.M. and D.L. Jones, "Efficient Detection in the Presence of Angular Spreading," ICASSP 2002, Orlando, Florida, May 13-16, 2002.

111. Baraniuk, R.G., C. Burns, B. Hendricks, G. Henry, A. Hero, D. Johnson, D.L. Jones, J. Kusuma, R. Nowak, J. Odegard, L.C. Potter, K. Ramchandran, "Connexions: DSP Education for a Networked World," *Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Orlando, Florida, May 13-16, 2002.

112. S.V. Adve, A.F. Harris, C.J. Hughes, D.L. Jones, R.H. Kravets, K. Nahrstedt, D.G. Sachs, R. Sasanka, J. Srinivasan, and W. Yuan, "The Illinois GRACE Project: Global Resource Adaptation through Cooperation," Proc. Workshop on Self-Healing, Adaptive, and Self-Managed Systems, New York, June 23, 2002.

113. Iyer, N., Kramer, M., Larsen, J., Lansing, C., Bilger, R., Jones, D., O'Brien, W., Wheeler, B., and Feng, A. (2002). Evaluating the use of objective metrics for assessing a binaural signal processing algorithm.International Hearing Aid Research Conference, Lake Tahoe, CA, August 2002.

114. Krongold, B.S. and D.L. Jones, "A Study of Active Constellation Extension for PAR Reduction in OFDM," *Proceedings of the 7th International OFDM Workshop*, Hamburg, Germany, Sept 2002, pp. 107–111.

115. Yuan, W., K. Nahrstedt, S. Adve, D.L. Jones, and R. Kravets, "Design and Evaluation of A Cross-Layer Adaptation Framework for Mobile Multimedia Systems," Proc. of SPIE/ACM Multimedia Computing and Networking Conference (MMCN'03), Santa Clara, CA, January 2003, pp. 1–13.

116. Roberts, D.S. and D.L. Jones, "Consistent Time-Frequency Representations," *Proc. International Conference on Acoustics, Speech, and Signal Processing*, Hong Kong, April 6–10, 2003.

117. Krongold, B.S. and D.L. Jones, "PAR Reduction in OFDM Via Active Constellation Extension," *Proc. International Conference on Acoustics, Speech, and Signal Processing*, Hong Kong, April 6–10, 2003.

118. Chou, J., D.G. Sachs, K. Ramchandran, and D.L. Jones, "Audio Data Hiding with Application to Surround Sound," *Proc. International Conference on Acoustics, Speech, and Signal Processing*, Hong Kong, April 6–10, 2003.

119. Appadwedula, S., R.G. Baraniuk, M.J. Berry, M. Butala, H. Choi, M. Haun, D.L. Jones, M.L. Kramer, D. Moussa, L.C. Potter, D.G. Sachs, B. Wade, and R. Wagner, "Open-Content Signal Processing Laboratories in Connexions," *Proc. International Conference on Acoustics, Speech, and Signal Processing*, Hong Kong, April 6–10, 2003.

120 Jones, D.L., "Four-Dimensional Sound Source Recovery From Arbitrary Acoustic Arrays," *Proc. IEEE International Conference on Multimedia and Expo*, Baltimore, Maryland, July 6–9, 2003.

121 Engel, J., J. Chen, X. Wang, Z. Fan, C. Liu, and D.L. Jones, "Technology Development of Integrated Multi-modal and Flexible Tactile Skin for Robotics Applications," *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Las Vegas, NV, 2003.

122. Atkinson, I., F. Kamalabadi, D.L. Jones, and M. Do, "Adaptive Wavelet Thresholding for Multichannel Signal Estimation," *Wavelets X Conference*, Proc. SPIE vol. 5207, San Diego, CA, August 4–8, 2003.

123. Sachs, D.G., S. Adve, and D.L. Jones, "Cross-Layer Adaptive Video Coding to Reduce Energy on General-Purpose Processors," *Proc. IEEE International Conference on Image Processing*, Barcelona, Spain, Sept. 14–17, 2003.

124. Atkinson, I., F. Kamalabadi, S. Mohan, and D.L. Jones, "Wavelet-Based 2-D Multichannel Signal Estimation," *Proc. IEEE International Conference on Image Processing*, Barcelona, Spain, Sept. 14–17, 2003.

125. Atkinson, I., F. Kamalabadi, and D.L. Jones, "Wavelet-Based Hyperspectral Image Estimation," *Proc. IGARSS conference*, Toulouse, France, 2003.

126. Mohan, S., M.L. Kramer, B.C. Wheeler, and D.L. Jones, "Localization of nonstationary sources using a coherence test," *Proc. IEEE Workshop on Statistical Signal Processing*, St. Louis, Missouri, September 2003.

127. Haun, M.A., D.L. Jones, and W.D. O'Brien, Jr., "Aberration Correction with the Overdetermined Fan-Filtering (OFF) Algorithm," *Proc. 2003 IEEE International Ultrasonics Symposium*, Honolulu, Hawaii, Oct. 6-8 2003.

128. Lockwood, M. E., Jones, D. L., Su, Q., and Miles, R. N., "Beamforming with collocated microphone arrays," *Journal of the Acoustical Society of America*, vol. 114, no. 4, pt. 2, October 2003, p. 2451.

129. Mohan, S., M.E. Lockwood, D.L. Jones, Q. Su, and R. N. Miles, "Sound source localization with a gradient array using a coherence test," *Journal of the Acoustical Society of America*, vol. 114, no. 4, pt. 2, October 2003, p. 2451.

130. Jones, D.L., "Property-Restoral Algorithms for Blind Equalization of OFDM," *Proc. 37th Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, November, 2003, pp. –.

131. Lockwood, M.E., D.L. Jones, C.R. Lansing, W.D. O'Brien, Jr., B.C. Wheeler, A.S. Feng, "Effect of Multiple Nonstationary Sources on MVDR Beamformers," *Proc. 37th Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, November, 2003, pp. –.

132. Zhang, S.-, P.K. Wong, D.G. Sachs, R. Koetter, and D.L. Jones, "Stability Analysis of Stochastic Sensor Networks," *Proc. 37th Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, November, 2003, pp. –.

133. Choi, H., R.G. Baraniuk, D.B. Williams, T. Barnwell, and D.L. Jones, "Open-Content Signal Processing Labs in Connexions," *Texas Instruments Developers' Conference*, Houston, TX, February 18-20, 2004.

134. Jones, D., M.E. Lockwood, C.R. Lansing, and A.S. Feng, "Improved BTE hearing aid directivity using a directional microphone array," *Journal of the Acoustical Society of America*, vol. 115, no. 5, Pt. 2, May 2004, p. 2598.

135. Appadwedula, S., R.G. Baraniuk, M.J. Berry, M. Butala, H. Choi, M.A. Haun, D.L. Jones, M.L. Kramer, D. Moussa, L.C. Potter, D.G. Sachs, B. Wade, R. Wagner, "Open-Content Signal Processing Labs in Connexions," *American Society of Engineering Education Annual Conference*, Salt Lake City, June 20-23, 2004.

136. Schmitz, C., N. Iyer, M.E. Lockwood, C.R. Lansing, and D.L. Jones, "Enhanced BTE Directivity Using a Directional Microphone Array," *IHCON 2004: International Hearing Aid Research Conference*, Tahoe City, CA, August 25–29, 2004.

137. Zhang, S.-C., I. Koprulu, R. Koetter, and D.L. Jones, "Feasibility Analysis of Stochastic Sensor Networks," *Proc. 1st IEEE Comm. Soc. Conf. on Sensor and Ad Hoc Communications and Networks*, Pacific Grove, CA, October 4–7, 2004.

138. Krongold, B.S., K. Ramchandran, and D.L. Jones, "Optimal Power and Rate Allocation in OFDM Systems," *Proceedings of the International Conference on Optimization Techniques and Applications*, Ballarat, Victoria, Australia, Dec 9-11, 2004.

139. Woo, G. and D.L. Jones, "Peak Power Reduction in MIMO OFDM via Active Channel Extension," *IEEE International Conference on Communi cations*, Seoul, Korea, May 15-19, 2005.

Electronic publications

1. M.J. Berry, M.L. Kramer, S. Appadwedula, D.G. Sachs, D. Moussa, M.A. Haun, B. Wade, J.W. Janovetz, D.L. Jones, "DSP Laboratory with TI TMS320C54x," `http://cnx.rice.edu/content/co` 2002.

2. J.P. Frantz, E. Yamai, K. Wang, K. Kriengchaipruck, M.J. Berry, M.L. Kramer, S. Appadwedula, D.G. Sachs, D. Moussa, M.A. Haun, B. Wade, J.W. Janovetz, D.L. Jones, "DSP Laboratory with TI TMS320C54x, (International Demo)," `http://cnx.rice.edu/content/col10227/latest,` 2004.

3. M.J. Berry, M.L. Kramer, S. Appadwedula, D.G. Sachs, D. Moussa, M.A. Haun, B. Wade, M. Butala, M. Kleffner, M. Frutiger, R. Morrison, J.N. Laska, J.W. Janovetz, A. Kulothungun, R. Cantzler, D.L. Jones, "Digital Signal Processing Laboratory," `http://cnx.rice.edu/content/col10236/` 2004.

4. D.L. Jones, "Digital Filter Structures and Quantization Error Analysis," `http://cnx.rice.edu/content/` 2005.

5. D.L. Jones, "Filter Structures," `http://cnx.rice.edu/content/m11917/latest/`, 2004.

6. D.L. Jones, "FIR Filter Structures," `http://cnx.rice.edu/content/m11918/latest/`, 2004.

7. D.L. Jones, "IIR Filter Structures," `http://cnx.rice.edu/content/m11919/latest/`, 2004.

8. D.L. Jones, "State-Variable Representation of Discrete-Time Systems," `http://cnx.rice.edu/content/m` 2004.

9. D.L. Jones, "Fixed-Point Number Representation," `http://cnx.rice.edu/content/m11930/latest/`, 2004.

10. D.L. Jones, "Fixed-Point Quantization," `http://cnx.rice.edu/content/m11921/latest/`, 2004.

11. D.L. Jones, "Finite-Precision Error Analysis," `http://cnx.rice.edu/content/m11922/latest/`, 2004.

12. D.L. Jones, "Input Quantization Noise Analysis," `http://cnx.rice.edu/content/m11923/latest/`, 2004.

13. D.L. Jones, "Quantization Error in FIR Filters," `http://cnx.rice.edu/content/m11924/latest/`, 2004.

14. D.L. Jones, "Data Quantization in IIR Filters," `http://cnx.rice.edu/content/m11925/latest/`, 2004.

15. D.L. Jones, "IIR Coefficient Quantization Analysis," `http://cnx.rice.edu/content/m11926/latest/`, 2004.

16. D.L. Jones, "Scaling," `http://cnx.rice.edu/content/m11927/latest/`, 2004.

17. D.L. Jones, "Limit Cycles," `http://cnx.rice.edu/content/m11928/latest/`, 2004.

Patents

1. Jones, D.L. and B.J. Hunsinger, "Apparatus and Method for Wide Bandwidth Adaptive Filtering," *Assignee: Board of Trustees of the University of Illinois*, U.S. Patent No. 5,272,663, issued December 21, 1993.

2. Jezwinski, S., and D.L. Jones, "Method and Apparatus for Testing Fiber Optic Telephone Lines," *Assignee: Teradyne Corp.*, U.S. Patent No. 5621517, issued April 15, 1997.

3. Krongold, B.S., K. Ramchandran, and D.L. Jones, "Section Division Operating Point Determination Method for Multicarrier Communication Systems," *Assignee: Board of Trustees of the University of Illinois*, U.S. Patent No. 6,400,773, issued June 4, 2002.

4. Feng, A.S., Liu, C., Bilger, R.C., Jones, D.L., Lansing, C.R., O'Brien, W.D. and Wheeler, B.C., "Binaural signal processing techniques," *Assignee: Board of Trustees of the University of Illinois*, Australian Patent No. 748113 (WO 00/30404), 2002.

5. Feng, A.S., Liu, C., Bilger, R.C., Jones, D.L., Lansing, C.R., O'Brien, W.D. and Wheeler, B.C., "Dual Channel Extraction Algorithm/System for High-Fidelity Signal Localization, Characterization and interference Suppression," Invention Disclosure: UIUC Technology #97060, September 19, 1997; filed on November 16, 1998, (US Serial number 09/193,058).

6. Jones, D.L., Lockwood, M.E., Bilger, R.C., Feng, A.S., Lansing, C.R., O'Brien, W.D., Wheeler, B.C., Elledge, M. and Liu, C., "Interference suppression techniques," (UIUC Tech ID - TF00015) - Patent application filed in May 10, 2000 (US Serial number 09/568,430).

7. Feng, A.S., Jones, D.L., B.C. Wheeler, R.C. Bilger, C.R. Lansing, and W.D. O'Brien, "Intrabody Communication for a Hearing Aid," *Assignee: Board of Trustees of the University of Illinois*, Application No. 09/568,435, filed on May 10, 2000.

8. Jones, D.L., Lockwood, M.E., Bilger, R.C., Feng, A.S., Lansing, C.R., O'Brien, W.D., Wheeler, B.C., Elledge, M. and Liu, C. Interference suppression techniques. PCT patent application (on the basis of US patent application serial number 09/568,430) filed on May 10, 2001 (WO 01/87011 A2).

9. Jones, D.L., C. Malick, X. Qi, M. Parikh, S.J. Franke, J.B. Larsen, and C.D. Schmitz, "Antenna Placement on In-The-Ear (ITE), Completely-In-The Ear (CIC), and Behind-The-Ear (BTE) Hearing Aid Shells for Bi-Direc tional Communication Around the Head," *Assignee: Board of Trustees of the University of Illinois*, Illinois IP# TF02129, US Patent Application No. , filed on January 21, 2003.

10. A.S. Feng, M.E. Lockwood, D.L. Jones, R.C. Bilger, W.D. O'Brien, and B.C. Wheeler, "System and Methods for Interfe rence Suppression with Directional Sensing Patterns," *Assignee: Board of Trustees of the University of Illinois*, US Patent Application No. 10/409,969, filed April 9, 2003. (Illinois IP no. TF02007)

11. A.S. Feng, M.E. Lockwood, D.L. Jones, R.C. Bilger, W.D. O'Brien, and B.C. Wheeler, "System and Methods for Interference Suppression with Directional Sensing Patterns," *Assignee: Board of Trustees of the University of Illinois*, PCT Patent Application, based on US Patent Application No. 10/4 09,969, filed April 6, 2004. (Illinois IP no. TF02007)

12. Ratnam, R., A.S. Feng, D.L. Jones, C.R. Lansing, and W.D. O'Brien, "Reverberation Time Evaluation," *Assignee: Board of Trustees of the University of Illinois*, Illinois IP# TF03008, US Patent Application Serial No. 60/466,133, filed 2003 or 2004.

13. Nollett, B.S., J.P. Oliver, and D.L. Jones, "Modulation Schemes for Reduced Power Amplifier Backoff," *Assignees: Motorola, Inc. and the Board of Trustees of the University of Illinois*, US Patent Application Serial No. 10/841,657, filed May 7, 2004.

Software

1. Jones, D.L. and R.G. Baraniuk, "aok4.c — an adaptive-kernel time-frequency analysis tool," `http://jazz.rice.edu/software/TFA/AOK/index.html`, February 1996.

Reviews, Sidebars, Columns

1. Jones, D.L., "Review of the MONARCH software package," *IEEE Spectrum*, vol. 26, no. 7, July 1989, p. 14.
2. D.G. Sachs, W. Yuan, C.J. Hughes, A.F. Harris III, S.V. Adve, Jones, D.L., R.H. Kravets, and K. Nahrstedt, "GRACE: A Cross-Layer Adaptation Framework for Saving Energy," *IEEE Computer*, vol. 36, no. 12, December 2003, pp. 56–57.

Theses

1. Jones, D.L., "A Study of Windows for the Short-Time Fourier Transform," *M.S.E.E. Thesis, Rice University*, Houston, Texas, November 1985.
2. Jones, D.L., "A High Resolution Data Adaptive Time-Frequency Representation," *Ph.D. Dissertation, Rice University*, Houston, Texas, May 1987.
3. Harris, W.M., "Computationally Efficient Frequency Rate Estimation Algorithms," *M.S.E.E. Thesis, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, December 1989.
4. Koczwara, T.E., "On Mask Selection for Time-Varying Filtering Using the Wigner Distribution," *M.S.E.E. Thesis, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, January 1990.
5. Raitz, J.E., "An Image Compression Card for the Macintosh II," *M.S.E.E. Thesis, University of Illinois at Urbana-Champaign*, D.L. Jones, co-advisor, 1990.
6. McBride, S.T., "An Efficient Polynomial Phase Estimator," *M.S.E.E. Thesis, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, June 1991.
7. Miyazu, K., "Adaptive FIR Filtering Using Acoustic Charge Transport Device with Bypass Digital Filter," *M.S.E.E. Thesis, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, March 1992.
8. Baraniuk, R.G., "Shear Madness: Signal-Dependent and Metaplectic Time-Frequency Representations, *Ph.D. Dissertation, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, August 1992.

9. Young, C., "High-Peformance Digital Filters," *Ph.D. Dissertation, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, September 1992.

10. Czerwinski, R.N., "Adaptive Time-Frequency Analysis Using a Cone-Shaped Kernel," *M.S.E.E. Thesis, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, December 1992.

11. Tziortzis, A., "Image Spectrum Extrapolation in Magnetic Resonance Imaging," *M.S.E.E. Thesis, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, January 1993.

12. Canfield, K.G., "An Adaptive, Hyperbolic-Class Time-Frequency Representation," *M.S.E.E. Thesis, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, April 1993.

13. Sayeed, A.M., "Optimal Kernels for Nonstationary Spectral Estimation," *M.S.E.E. Thesis, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, September 1993.

14. Yu, G.S., "Modifications to the NLMS Algorithm When Used for Acoustic Echo Cancellation," *M.S.E.E. Thesis, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, September 1994.

15. Siampos, G.S. "Least Mean Squares Adaptive Filtering Using Acoustic Charge Transport Technology", *M.S.E.E. Thesis, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, December 1994.

16. Kramer, M.L., "Adaptive Time-Frequency Analysis and Synthesis," *M.S.E.E. Thesis, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, December 1994.

17. Ren, Y., "Adaptive Bases for Solving Partial Differential Equations", *M.S.E.E. Thesis, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, August 1995.

18. Sayeed, A.M., "Statistical Time-Frequency Analysis," *Ph.D. Dissertation, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, August 1996.

19. Czerwinski, R.N., "Boundary Detection in Ultrasonic Speckle," *Ph.D. Dissertation, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, October 1996.

20. Czerwonka, G.J., "Instantaneous Frequency Analyses for Mechanical Equipment Fault Identification," *M.S.E.E. Thesis, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, June 1997.

21. Steward, D.L., "Kalman Filter Analysis of an Animal Head-Motion Estimation System," *M.S.E.E. Thesis, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, June 1997.

22. Yang, V.Y., "A Vector Constant Modulus Algorithm for Shaped Constellation Equalization," *M.S.E.E. Thesis, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, August 1997.

23. Krongold, B.S., "Power and Bandwidth Optimization for Multicarrier Communication Systems," *M.S.E.E. Thesis, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, September 1997.

24. Rao, A.M., "Optimal Quadratic Array Detection Using Quadratic Time-Frequency and Time-Scale Representations," *M.S.E.E. Thesis, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, September 1997.

25. Kwok, H. K.-C., "Improved Instantaneous Frequency Estimation using Adaptive Short-Time Fourier Transform," *M.S.E.E. Thesis, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, May 1998.

26. Schwartz, B., "Quadratic and Instantaneous Frequency Analysis of Helicopter Gearbox Faults," *M.S.E.E. Thesis, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, July 1998.

27. Qian, L., "Joint Source-Channel Matching for Wireless Video Transmission," *M.S.E.E. Thesis, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, November 1998.

28. Appadwedula, S., "Joint Source-Channel Matching for Wireless Image Transmission," *M.S.E.E. Thesis, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, December 1998.

29. Haun, M.A., "The Fractionally Spaced Vector Constant Modulus Algorithm," *M.S.E.E. Thesis, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, August 1999.

30. Lockwood, M.E., "Development and Testing of a Frequency Domain Minimum Variance Algorithm for use in a Binaural Hearing Aid," *M.S.E.E. Thesis, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, 1999.

31. Janovetz, J.W., "A Flexible Digital Communications System for Education," *M.S.E.E. Thesis, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, September 1999.

32. Elledge, M.E, "Real-Time Implementation of a Frequency-Domain Array Processor," *M.S.E.E. Thesis, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, May 2000.

33. Burke, B., "Subband Channelized Radar Detection and Bandwidth Estimation for FPGA Implementation," *M.S.E.E. Thesis, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, May 2000.

34. Sachs, D.G, "Wireless Image Transmission over Fading Channels Using Multiple-Description Concatenated Codes," *M.S.E.E. Thesis, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, May 2000.

35. Yang, K.L., "Comparative Analysis of Binaural Acoustic Processing Algorithms," *M.S.E.E. Thesis, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, May 2000.

36. Kwok, H. K.-C., "Shape Up: Peak Power Reduction via Constellation Shaping," *Ph.D. Dissertation, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, December 2000.

37. Rao, A.M., "Efficient Array Processing for Partial Signal Coherence," *Ph.D. Dissertation, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, April 2001.

38. Qian, L., "Joint Source-Channel Video Transmission," *Ph.D. Dissertation, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, September 2001.

39. Kramer, M.L., "Denoising via Empirical Bayesian Pursuit," *Ph.D. Dissertation, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, December 2001.

40. Krongold, B.S., "New Techniques for Multicarrier Communication Systems," *Ph.D. Dissertation, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, December 2001.

41. Roberts, D.S., "Applications and Extensions of Quadratic Signal Representations," *Ph.D. Dissertation, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, September 2002.

42. Nollett, B., "Efficient Nonlinear Adaptive Filtering Using the Wiener-Hammerstein System Model," *M.S.E.E. Thesis, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, October 2002.

43. Wade, B., "Low-Power Transceiver Architectures for an Intelligent Hearing Aid," *M.S.E.E. Thesis, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, November 2002.

44. Berry, M., "Improving Interference Attenuation of Wireless Hearing-Aid Beamformers by Increasing Number of Sensors," *M.S.E.E. Thesis, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, January 2003.

45. Haun, M.A., "New Approaches to Aberration Correction in Medical Ultrasound Imaging," *Ph.D. Dissertation, University of Illinois at Urbana-Champaign*, D.L. Jones and W.D. O'Brien, co-advisors, May 2003.

46. Appadwedula, S., "Energy-Efficient Sensor Networks for Detection Applications," *Ph.D. Dissertation, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, May 2003.

47. Detwiler, T., "Active constellation extension compatible OFDM receiver designs," *M.S.E.E. Thesis, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, May 2003.

48. Mohan, S., "Localization of nonstationary sources with small directional arrays," *M.S.E.E. Thesis, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, August 2003.

49. Patel, P., "Blind Estimation of Reverberation T ime from Noisy Acoustic Signals," *M.S.E.E. Thesis, University of Illinois at Urbana-Cham paign*, D.L. Jones, advisor, December 2003.

50. Ingale, S.S., "A Comparison of Training Methods for Hidden Markov Models," *M.S.E.E. Thesis, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, May 2004.

51. Wong, P.K., "A Communication Protocol for Stochastic Sensor Networks," *M.S.E.E. Thesis, University of Illinois at Urbana-Champaign*, D.L. Jones, advisor, May 2004.

52. Jordan, J.C., "Design and Implementation of a S tochastic Wireless Sensor Network," *M.S.E.E. Thesis, University of Illinois at Urbana-Champaign*, D.L. Jones, ECE Reader, August 2004.

92. Lockwood, M.E., M.E. Elledge, K.L. Yang, D.L. Jones, R.C. Bilger, A.S. Feng, J.B. Larsen, C.R. Lansing, and B.C. Wheeler, "Simulation and Real-Time Implementation of a New Optimum Beamforming Technique," *Proc. Intl. Hearing Aid Research Conference*, Lake Tahoe, Nevada, Aug 23–27, 2000, pp. 32-33.

93. Yang, K.L., Lockwood, M.E., Elledge, M.E. and D.L. Jones, "A Comparison of Beamforming Algorithms for Binaural Acoustic Processing," *Proceedings of the 9th IEEE Digital Signal Processing Workshop*, Hunt, Texas, October 15–18, 2000.

94. Kramer, M.L., M. Haun, S. Appadwedula, D.G. Sachs, and D.L. Jones, "Effective Use of Projects in DSP Laboratory Instruction," *Proceedings of the First IEEE Signal Processing Education Workshop*, Hunt, Texas, October 15–18, 2000.

95. Haun, M., D.L. Jones, and W.D. O'Brien, Jr., "Efficient Three-Dimensional Imaging from a Small Cylindrical Aperture," *Proc. of the IEEE Ultrasonics Symposium*, San Juan, Puerto Rico, October 22–25, 2000, pp. 1589–1592.

96. Larsen, J., M.E. Lockwood, M. Elledge, C.R. Lansing, R.C. Bilger, W.D. O'Brien, D.L. Jones, B.C. Wheeler, and A.S. Feng, "Performance of a frequency based minimum variance beamforming algorithm for hearing impaired and normal hearing listeners," *24th Annual ARO MidWinter Meeting*, St. Petersburg Beach, Florida, February, 4-8, 2001.

97. Roberts, D.S. and D.L. Jones, "A Four-Parameter Quadratic Distribution," *IEEE International Conf. on Acoustics, Speech, and Signal Processing*, Salt Lake City, Utah, May 7–11, 2001.

98. Jones, D.L., "Designing Effective DSP Laboratory Courses," *IEEE International Conf. on Acoustics, Speech, and Signal Processing*, Salt Lake City, Utah, May 7–11, 2001.

99. Sachs, D.G., I. Kozintsev, M. Yeung, and D.L. Jones, "Hybrid ARQ for robust video streaming over wireless LANs," *Proc. Information Technology: Coding and Computing*, Las Vegas, NV, 2001, pp. 317–321.

100. Haun, M.A., D.L. Jones, and W.D. O'Brien, Jr., "Efficient Three-Dimensional Cylindrical-Geometry Ultrasound Imaging," *Journal of the Acoustical Society of America*, vol. (presented at the Acoustical Society of America Spring 2001 Meeting, Chicago, IL, June 4–8, 2001).

101. Lockwood, M.E., Yang, K.L., Jones, D.L., Bilger, R.C., Lansing, C.R., Liu, C., O'Brien, W.D., Wheeler, B.C. and Feng, A.S. (2001) Robustness of beamforming algorithms with head-related transfer functions. J. Acoust. Soc. Am. 109, p. 2493.

102. Zheng, Y., Lockwood, M.E., Wheeler, B.C., Jones, D.L., Feng, A.S., O'Brien, W.D., Bilger, R.C., Lansing, C.R. (2001) Comparison of binaural beamformers for speech extraction in complex auditory scenes. J. Acoust. Soc. Am. 109.

103. Berry, M. Lockwood, M.E., Jones, D.L., Bilger, R.C., Lansing, C.R., Liu, C., O'Brien, W.D., Wheeler, B.C. and Feng, A.S. (2001) Beamformer improvement with increasing number of sensors. J. Acoust. Soc. Am. 109.

Patent Claims

1.-33. (Canceled).

34. (Previously presented) A method of signal processing, comprising:

(a) detecting an acoustic excitation at both a first location to provide a corresponding first signal and at a second location to provide a corresponding second signal, the excitation being a composite of a desired acoustic signal from a first source and an interfering acoustic signal from a second source spaced apart from the first source;

(b) determining location of the second source relative to the first source as a function of the first and second signals, which includes delaying each of the first and second signals by several time intervals to provide several delayed first signals and several delayed second signals and providing a time increment representative of separation of the first source from the second source; and

(c) generating a characteristic signal representative of the desired acoustic signal during performance of said determining, the characteristic signal being a function of the time increment.

35. (Previously presented) The method of claim 34, wherein the characteristic signal corresponds to spectral content of the desired acoustic signal and further comprising providing an output signal representative of the desired acoustic signal as a function of the characteristic signal.

36. (Canceled).

37. (Currently amended) The method of claim 34, wherein said determining includes:

establishing a signal pair, the signal pair having a first member from the delayed first signals and a second member from the delayed second signals, the characteristic signal being determined from the signal pair.

38. (Previously presented) The method of claim 34, further comprising providing an output signal representative of the desired acoustic signal, and wherein the desired acoustic signal includes speech and the output signal is provided by a hearing aid device.

39. (Currently amended) The method of claim 34, wherein said determining further includes:

(b1) converting the first and second signals from an analog representation to a discrete representation;

(b2) transforming the first and second signals from a time domain representation to a frequency domain representation; and

(b3) establishing a signal pair representative of separation of the first source from the second source, the signal pair having a first member from the delayed first signals and a second member from the delayed second signals.

40. (Currently amended) The method of claim 39, wherein the characteristic signal corresponds to a fraction with a numerator determined from at least the first and second members, and a denominator determined from at least the time increment.

41. (Previously presented) The method of claim 39, wherein said generating further includes:

(c1) determining the characteristic signal from the signal pair and the first time increment, the characteristic signal being representative of spectral content of the desired acoustic signal;

(c2) transforming the characteristic signal from a frequency domain representation to a time domain representation; and

(c3) providing an audio output signal representative of the desired acoustic signal as a function of the characteristic signal.

42. (Currently amended) The method of claim 41, further comprising establishing a further time increment corresponding to separation of the first source from the second source by comparing the delayed first and second signals, and

wherein the time increment corresponds to a first phase difference, the further time increment corresponds to a second phase difference, and the characteristic signal includes a spectral representation determined from at least the first and second phase differences.

43. (Canceled).

44. (Previously presented) The method of claim 34, wherein separation of the second source is within five degrees of the first source relative to a zero degree azimuthal

reference axis intersecting the first source and a midpoint situated between the first and second locations.

45. (Previously presented) The method of claim 34, further comprising;

(d) establishing a number of location signals each corresponding to a different location relative to the first source; and

(e) selecting the characteristic signal from the location signals, the characteristic signal being representative of the location of the second source relative to the first source, the characteristic signal including a spectral representation of the desired acoustic signal.

46. (Previously presented) A method of signal processing, comprising:

(a) detecting an acoustic excitation at a first location to provide a corresponding first signal and at a second location to provide a corresponding second signal, the excitation being a composite of a desired acoustic signal from a first source and an interfering acoustic signal from a second source spaced apart from the first source;

(b) localizing the second source relative to the first source as a function of the first and second signals, said localizing including establishing a number of location signals each corresponding to a different location relative to the first source, delaying each of the first and second signals by a number of time intervals to provide a number of delayed first signals and a number of delayed second signals, and establishing a signal pair that has a first member from the delayed first signals and a second member from the delayed second signals; and

(c) generating a characteristic signal from the location signals, wherein the characteristic signal includes a spectral representation of the desired acoustic signal from the first source, corresponds to position of the second source, and is determined from the signal pair.

47. (Previously presented) The method of claim 46, further comprising providing an output signal representative of the desired acoustic signal as a function of the characteristic signal.

48. (Currently amended) The method of claim 46, wherein said localizing includes:

determining a time increment representative of separation of the first source from the second source, the characteristic signal being a function of the time increment.

49. (Canceled).

50. (Previously presented) The method of claim 46, further comprising providing an output signal representative of the desired acoustic signal, and wherein the desired acoustic signal includes speech and the output signal is provided by a hearing aid device.

51. (Currently amended) The method of claim 46, wherein said localizing further includes:

(b1) converting the first and second signals from an analog representation to a discrete representation;

(b2) transforming the first and second signals from a time domain representation to a frequency domain representation; and

(b3) establishing a first time increment and a signal pair each representative of separation of the first source from the second source, the signal pair having a first member from the delayed first signals and a second member from the delayed second signals.

52. (Previously presented) The method of claim 51, wherein the characteristic signal corresponds to a fraction with a numerator determined from at least the first and second members, and a denominator determined from at least the first time increment.

53. (Previously presented) The method of claim 51, wherein said generating further includes:

(c1) determining the characteristic signal from the signal pair and the first time increment;

(c2) transforming the characteristic signal from a frequency domain representation to a time domain representation; and

(c3) providing an audio output signal representative of the desired acoustic signal as a function of the characteristic signal.

54. (Previously presented) The method of claim 53, further comprising establishing a second time increment corresponding to separation of the first source from the second source by comparing the delayed first signals and delayed second signals, and

wherein the first time increment corresponds to a first phase difference, the second time increment corresponds to a second phase difference, and the spectral representation of the characteristic signal is determined from at least the first and second phase differences.

55. (Canceled).

56. (Previously presented) The method of claim 1, wherein separation of the second source is within five degrees of the first source relative to a zero degree azimuthal reference axis intersecting the first source and a midpoint situated between the first and second locations.

57. (New) The method of claim 34, wherein the characteristic signal corresponds to a fraction with a numerator determined from a difference between a first member of the delayed first signals and a second member of the delayed second signals, and a denominator determined from at least the time increment.

58. (New) The method of claim 57, which includes providing the delayed first signals from a first multistage delay line and the delayed second signals from a second multistage delay line, the first member being output by a stage of the first delay line corresponding to the location of the second source and the second member being output by a stage of the second delay line corresponding to the location of the second source, and a different stage

of each of the first delay line and the second delay line corresponding to location of the first source.

59. (New) The method of claim 58, wherein the difference is representative of a minimized interfering acoustic signal level and provides the characteristic signal representative of spectral content of the desired acoustic signal.

60. (New) The method of claim 46, wherein the generating includes determining the characteristic signal as a fraction with a numerator being a function of a difference between one of the delayed first signals and one of the delayed second signals, the difference being representative of a minimized interfering acoustic signal level, and the fraction having a denominator determined as a function of at least the first time increment.

61. (New) A method of signal processing, comprising:

detecting an acoustic excitation at both a first location to provide a corresponding first signal and at a second location to provide a corresponding second signal, the excitation being a composite of a desired acoustic signal from a first source and an interfering acoustic signal from a second source spaced apart from the first source;

incrementally delaying the first signal to provide a number of delayed first signals and the second signal to provide a number of delayed second signals, a number of different pairings of the delayed first signals and the delayed second signals representing different locations;

localizing the second source relative to one of the different locations as a function of a difference between the members of a corresponding one of the different pairings; and

generating a characteristic signal representative of spectral content of the desired acoustic signal based on the difference and a time increment corresponding to distance separating the first source and the second source.

62. (New) A method of signal processing, comprising:

detecting an acoustic excitation at both a first location to provide a corresponding first signal and at a second location to provide a corresponding second signal, the excitation being a composite of a desired acoustic signal from a first source and an interfering acoustic signal from a second source spaced apart from the first source;

selecting the desired acoustic signal by positioning a reference axis relative to the first source;

localizing the second source relative to the reference axis as a function of the first and second signals; and

generating a characteristic signal representative of the desired acoustic signal during performance of said localizing.

63. (New) The method of claim 62, which includes:

defining the reference axis relative to the first location and the second location; and

moving the reference axis to select a different acoustic signal.

64. (New) The method of claim 63, wherein the detecting the acoustic excitation is performed with a first sensor at the first location and a second sensor at the second location.

65. (New) The method of claim 63, wherein the method is performed with a hearing aid.

66. (New) The method of claim 63, wherein:

the localizing includes establishing a number of delayed first signals each corresponding to a different one of a number of first delay stages of a first delay line and a number of delayed second signals each corresponding to a different one of a number of second delay stages of a second delay line; and

the generating includes determining the characteristic signal as a function of a fraction with a numerator corresponding to a difference between one output of the first delay stages and one output of the second delay stages and a denominator corresponding to a time increment representative of a distance separating the first source and the second source.

US006002776A

United States Patent [19]

Bhadkamkar et al.

[11] **Patent Number: 6,002,776**

[45] **Date of Patent: *Dec. 14, 1999**

[54] **DIRECTIONAL ACOUSTIC SIGNAL PROCESSOR AND METHOD THEREFOR**

[75] Inventors: **Neal Ashok Bhadkamkar**, Palo Alto; **John-Thomas Calderon Ngo**, Sunnyvale, both of Calif.

[73] Assignee: **Interval Research Corporation**, Palo Alto, Calif.

[*] Notice: This patent issued on a continued prosecution application filed under 37 CFR 1.53(d), and is subject to the twenty year patent term provisions of 35 U.S.C. 154(a)(2).

[21] Appl. No.: **08/531,143**

[22] Filed: **Sep. 18, 1995**

[51] **Int. Cl.**[6] **H04B 3/20**

[52] **U.S. Cl.** **381/66**; 381/93; 379/410

[58] **Field of Search** 381/66, 94, 93, 381/96, 94.1, 94.7; 379/410, 390, 395, 406; 455/33.1

[56] **References Cited**

U.S. PATENT DOCUMENTS

3,500,000	3/1970	Kelly, Jr. et al. .	
4,544,926	10/1985	Giuli	343/18 E
4,750,147	6/1988	Roy et al. .	
4,751,738	6/1988	Widrow et al. .	
5,208,786	5/1993	Weinstein et al.	367/124
5,315,532	5/1994	Comon .	
5,371,789	12/1994	Hirano	379/410
5,383,164	1/1995	Sejnowski et al.	455/33.1
5,465,302	11/1995	Lazzari et al. .	

OTHER PUBLICATIONS

Jeffery C. Allen, "Advanced Beamforming Concepts: Source Localization Using the Bispectrum, Gabor Transform, Wigner–Ville Distribution, and Nonstationary Signal Representations," 25th Asilomar Conference on Signals, Systems, and Computers, vol. 2, 818–824, 1991.

Yeheskel Bar–Ness, "Broadband Interference Cancelation Using a Bootstrapped Approach," ICASSP '93 vol. III,, 372–375, 1993.

Zaka U. Bhatti and Steven Bibyk, "Adaptive Analog Circuits to Perform Signal Separation of Multidimensional Sensor Signals," IEEE International Conference on Systems Engineering, 391–394, Dayton, OH, Aug. 1–3, 1991.

Jean–Francois Cardoso, "Source Separation Using Higher Order Moments," ICASSP '89, 2109–2112, 1989.

Jean–Francois Cardoso, "Super–Symmetric Decomposition of the Fourth–Order Cumulant Tensor. Blind Identification of More Sources than Sensors," ICASSP '91, vol. V, 3109–3112, 1991.

Jean–Francois Cardoso, Adel Belouchrani, and Beate Leheld, "A New Composite Criterion for Adaptive and Iterative Blind Source Separation," ICASSP '94, vol. IV, revised version, pp. 1–4.

(List continued on next page.)

Primary Examiner—Ping Lee

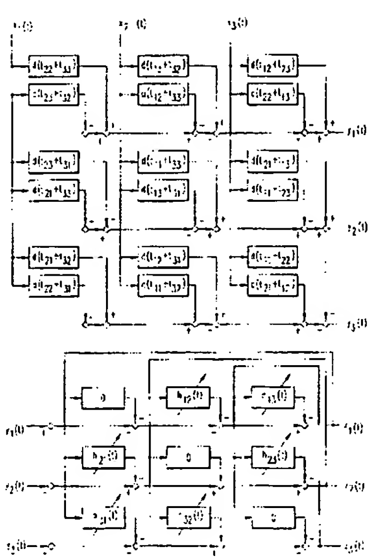
Attorney, Agent, or Firm—Limbach & Limbach L.L.P.

[57] **ABSTRACT**

Two or more microphones are mounted in an environment that contains an equal or lesser number of distinct sound sources. Acoustic energy from each source, with its attendant echoes and reverberation, impinges on each microphone. Using direction-of-arrival information, a first module attempts to extract the original source signals as if the acoustic environment were anechoic. Any residual crosstalk between the channels, which may be caused by echoes and reverberation, is removed by a second module. The first and second modules may be implemented using existing technology.

24 Claims, 3 Drawing Sheets

Microfiche Appendix Included
(1 Microfiche, 46 Pages)



OTHER PUBLICATIONS

Marc H. Cohen, "Analog VLSI Implementation of an Auto–Adaptive Synthetic Neutral Network for Real–Time Separation of Independent Signal Sources," Master's Thesis, Johns Hopkins University, May 1991.
Pierre Comon, Christian Jutten, and Jeanny Herault, "Blind Separation of Sources, Part II: Problems [sic] Statement," Signal Processing 24, 11–20, 1991.
Abdulkadir Dinc and Yeheskel Bar–Ness, A Forward/Backward Bootstrapped Structure for Blind Separation of Signals in a Multi–Channel Dispersive Environment, ICASSP '94, III, 376–379.
K. Farrell, RJ. Mammone, and J.L. Flanagan, "Beamforming Microphone Arrays for Speech Enhancement," ICASSP '92, vol. 1, 285–288, 1992.
J.L. Flanagan, J.D. Johnson, R. Zahn, and G.W. Elko, "Computer–Steered Microphone Arrays for Sound Transduction in Large Rooms," Journal of the Acoustical Society of America, 78(5), pp. 1508–1518, 1985.
O.L. Frost III, "An algorithm for linearly constrained adaptive array processing," Proceedings of the IEEE, 60(8), pp. 926–935, Aug. 1972.
L.J. Griffiths and C.W. Jim, "An alternative approach to linearly constrained adaptive beamforming," IEEE Transactions on Antennas and Propagation, AP–30(1), pp. 27–34, Jan. 1982.
C.W.K. Gritton and D.W. Lin, "Echo Cancellation Algorithms," IEEE ASSP Magazine, Apr. 30–38, 1984.
J.Herault and C. Jutten, "Space or Time Adaptive Signal Processing by Neutral Network Models", AIP Conference Proceedings, pp. 206–211, Snowbird, Utah 1986.
C. Jutten, L. Nguyen Thi, E. Djikstra, E. Vittoz, and J. Caelen, "Blind Separation of Sources: An Algorithm for Separation of Convolutive Mixtures," Higher Order Statistics, 275–278, 1992.
J.L. Lacoume and P. Ruiz, "Separation of Independent Sources of Correlated Inputs," IEEE Trans. Signal Processing 40(12), Dec. 1992.
Richard F. Lyon, "A Computational Model of Binaural Localization and Separation," ICASSP '83, 319–327, 1983.
David K. Mellinger, "Event Formation and Separation in Musical Sound," Ph.D. Dissertation, Stanford University, Dec. 1991.
O.M. Mracek Mitchell, Carolyn A. Ross, and G.H. Yates, "Signal Processing for a Cocktail Party Effect," Journal of the Acoustical Society of America 50(2), pp. 656–660, 1971.
L. Molgedey and H.G. Schuster, "Separation of a Mixture of Independent Signal Using Time Delayed Correlations," Physical Review Letters 72(23), 3634–3637, Jun. 6, 1994.
Tatsuya Morita, "Separation of Mixed Voices by Acoustic Parameter Optimization," Electronics and Communications in Japan, Part 3: Fundamental Electronic Science 74(5), 30–38, 1991.
Montse Najar, Miguel A. Lagunas, and Ignasi Bonet, "Blind Wideband Source Separation," ICASSP '94, vol. IV, 65–68, 1994.
J. Naylor and J. Porter, "An Effective Speech Separation System Which Requires No A Priori Information," ICASSP '91, vol. II, 937–940, 1991.
Harry F. Olson, "Directional Microphones", Journal of the Audio Engineering Society, pp. 190–194, Oct., 1967.
Thomas W. Parsons, "Separation of Speech from Interferring Speech by Means of Harmonic Selection," J. Acoust. Soc. Am. 60(4), 911–918, 1976.
John C. Platt and Federico Faggin, "Networks for the Separation of Sources that are Superimposed and Delayed," NIPS '91, pp. 730–737, 1992.
Richard Roy, "ESPRIT—Estimation of Signal Parameters via Rotational Invariance Techniques," Ph.D. dissertation, Stanford University, Stanford, CA 1987.
R.O. Schmidt, "A Signal Subspace Approach to Multiple Emitter Location and Spectral Estimation," Ph.D. Dissertation, Stanford University, Stanford, CA 1981.
Hans Werner Strube, "Separation of Several Speakers Recorded by Two Microphones (Cocktail–Party Processing)," Signal Processing 3, 355–364, 1981.
Lang Tong, Ruey–wen Liu, Victor C. Soon, and Yih–Fang Huang, Indeterminacy and Identifiability of Blind Identification, IEEE Trans. Circuits and Systems 38(5), 499–509, May 1991.
Lang Tong, Yujiro Inouye, and Rueywen Liu, "Eigenstructure–Based Blind Identification if Independent Signals," ICASSP '91, 3329–3332, 1991.
Dirk Van Compernolle, "Adaptive Filter Structures for Enhancing Cocktail Party Speech from Multiple Microphone Recordings," 1989.
Dirk Van Compernolle, Weiya Ma, Fei Xie, and Marc Van Diest, "Speech Recognition in Noisy Environments with the Aid of Microphone Arrays," Speech Communication 9, 433–442, 1990.
Ch. von der Malsburg and W. Schneider, "A Neural Cocktail–Party Processor," Biological Cybernetics 54, 29–40, 1986.
Mitchel Weintraub, "The GRASP Sound Separation System," ICASSP '84, 18A.6.1–18A.6.4, 1984.
Product Spec. Sheet, "Bose Aviation Headset," Bose Corporation, Framingham, Mass.
Product Spec. Sheet, "Knowles CF 2949 Data Sheet," Knowles Electronics, Inc., Itasca, Ill.
Product Spec. Sheet, "NCT proACTIVE Series," Noise Cancellation Technologies, Inc., Stamford, Conn.
Product Spec. Sheet, "Sennheiser HDC 451 NoiseGard Mobile Headphone," Sennheiser Electronic Corporation, Old Lyme, Conn.
Jean–Francois Cardoso and Eric Moulines, "Minimum Contrast Estimation with Applications to Array Processing," ICASSP '93, vol. IV, 384–387, 1993.
Yves Grenier, "A Microphone Array for Car Environments," Speech Communication 12, 25–39, 1993.
J.J. Hopfield, "Olfactory Computation and Object Perception," Proc. Natl. Acad. Sci. USA 88, 6462–6466, Aug. 1991.
Yutaka Kaneda, "Sound Source Localization for Wide–Band Signals Under a Reverberant Condition," J. Acoust. Soc. Japan (E), 14(1), 47–48, 1993.
Miguel A. Lagunas and Alba Pages, "Multitone Tracking with Coupled EKFs and High Order Learning," ICASSP '92, vol. V, 153–156, 1992.
F. Laichi, T. Aboulnasr, and W. Steenaart, "A New Efficient Implementation of the AIFIR Echo Tail Canceller for ISDN Application," ICASSP '91, 1545–1548, 1991.
Eric Moulines and Jean–Francois Cardoso, "Direction Finding Algorithms Using Fourth Order Statistics. Asymptotic Performance Analysis," ICASSP '92, vol. II, 437–440, 1992.
Montse Najar, Miguel A. Lagunas, and Ana I. Perez–Neira, "Source Separation Based on Coupled Single DOA Estimation Processors," ICASSP '93, vol. IV, 336–339, 1993.

Patrick M. Peterson, Nathaniel I. Durlach, William M. Rabinowitz, and Patrick M. Zurek, "Multimicrophone Adaptive Beamforming for Interference Reduction in Hearing Aids," Journal of Rehabilitation Research and Development 24(4), 103–110, Fall 1987.

Richard Roy and Thomas Kailath, "ESPRIT—Estimation of Signal Parameters via Rotational Invariance Techniques," Optical Engineering 29(4), 296–313, Apr. 1990.

Richard Roy and Thomas Kailath, "ESPRIT—Estimation of Signal Parameters Via Rotational Invariance Techniques," IEEE Trans. Acoust., Speech and Signal Processing 37(7), 984–995, Jul. 1989.

Harvey F. Silverman and Stuart E. Kirtman, "A Two–Stage Algorithm for Determining Talker Location from Linear Microphone Array Data," Computer Speech and Language 6, 129–152, 1992.

Richard M. Stern, Jr. and H. Steven Colburn, "Theory of Binaural Interaction Based on Auditory–Nerve Data. IV. A Model for Subjective Lateral Position," J. Acoust. Soc. Am. 64(1), Jul. 1978.

Noboru Sugie, Jie Huang, and Noboru Ohnishi, "Localizing Sound Source by Incorporating Biological Auditory Mechanism," IEEE Int'l Conf. on Neural Networks, II–243–250, 1988.

W.M. Wagenaars, "Localization of Sound in a Room with Reflecting Walls," J. Audio Eng. Soc. 38(3), 99–110, Mar. 1990.

Leland B. Jackson, "Digital Filters and Signal Processing", p. 146, 1986.

Gordon E. Peterson, "Session F. Speech Intelligibility and Perception", J. Acoust. Society, 32(7), Jul. 1960, p. 918.

Mitchell Weintraub "A Theory and Computational Model of Auditory Monaural Sound Separation", Ph.D. Dissertation, Department of Electrical Engineering, Stanford University, Aug. 1985.

Mitchel Weintraub "Computational Model for Separating Two Simultaneous Talkers," ICASSP '86, pp. 81–84, 1986.

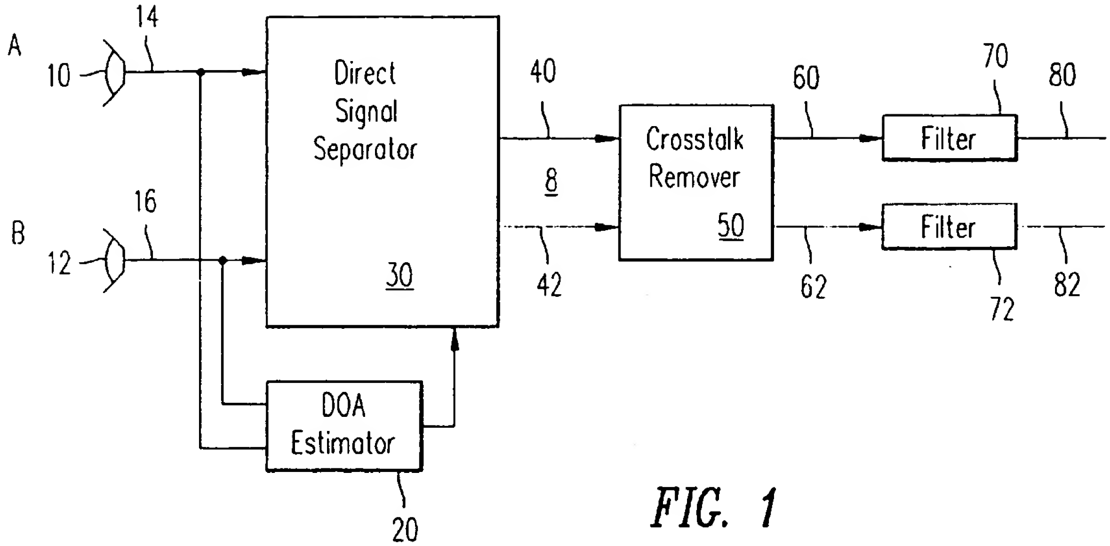


FIG. 1

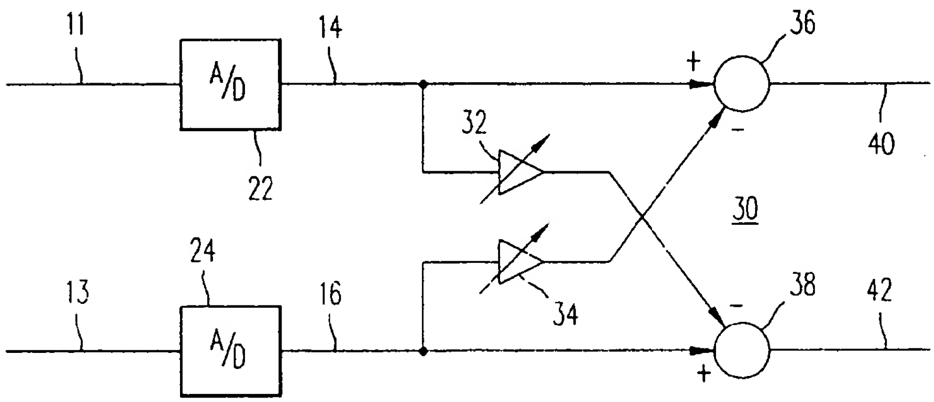


FIG. 2

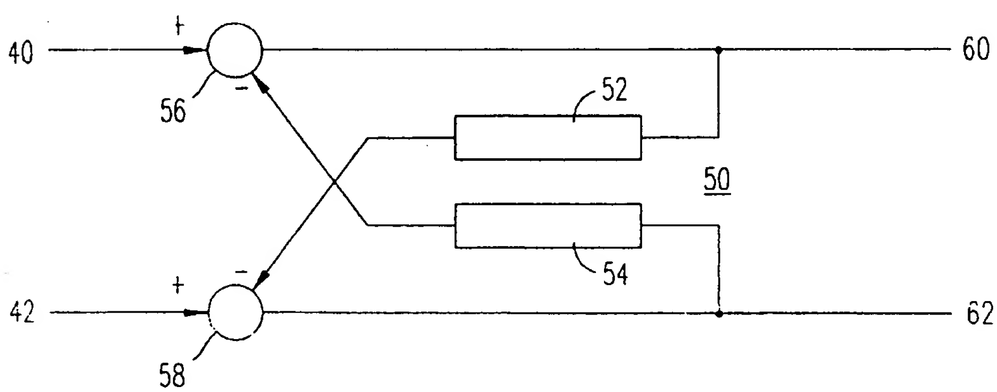


FIG. 3

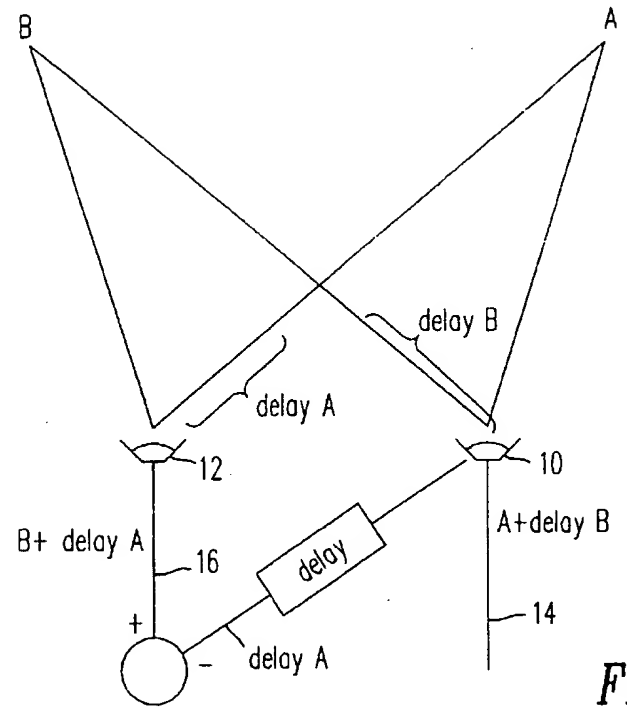

B
A
delay B
delay A
12
10
B+ delay A
A+delay B
16
delay
14
+
−
delay A

FIG. 4

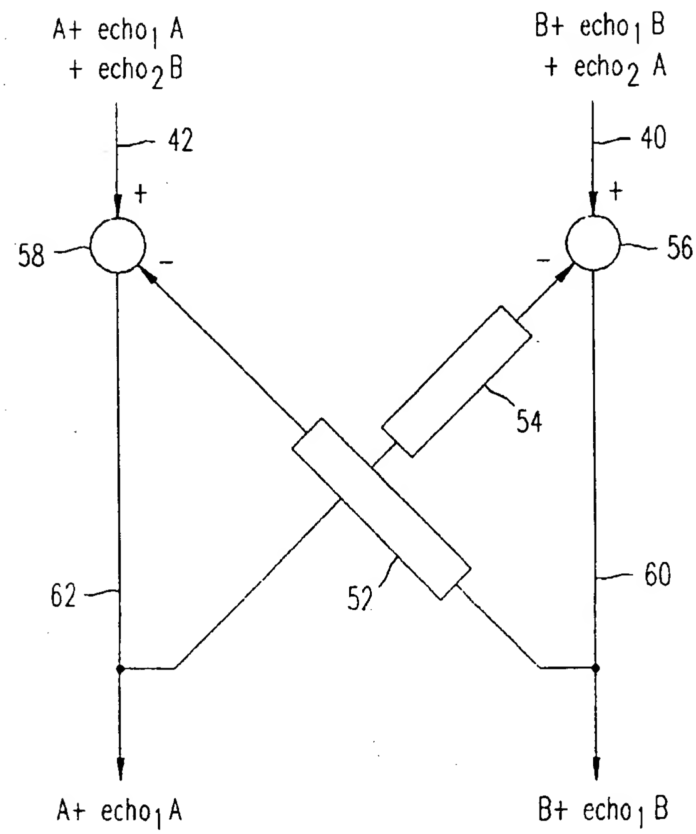
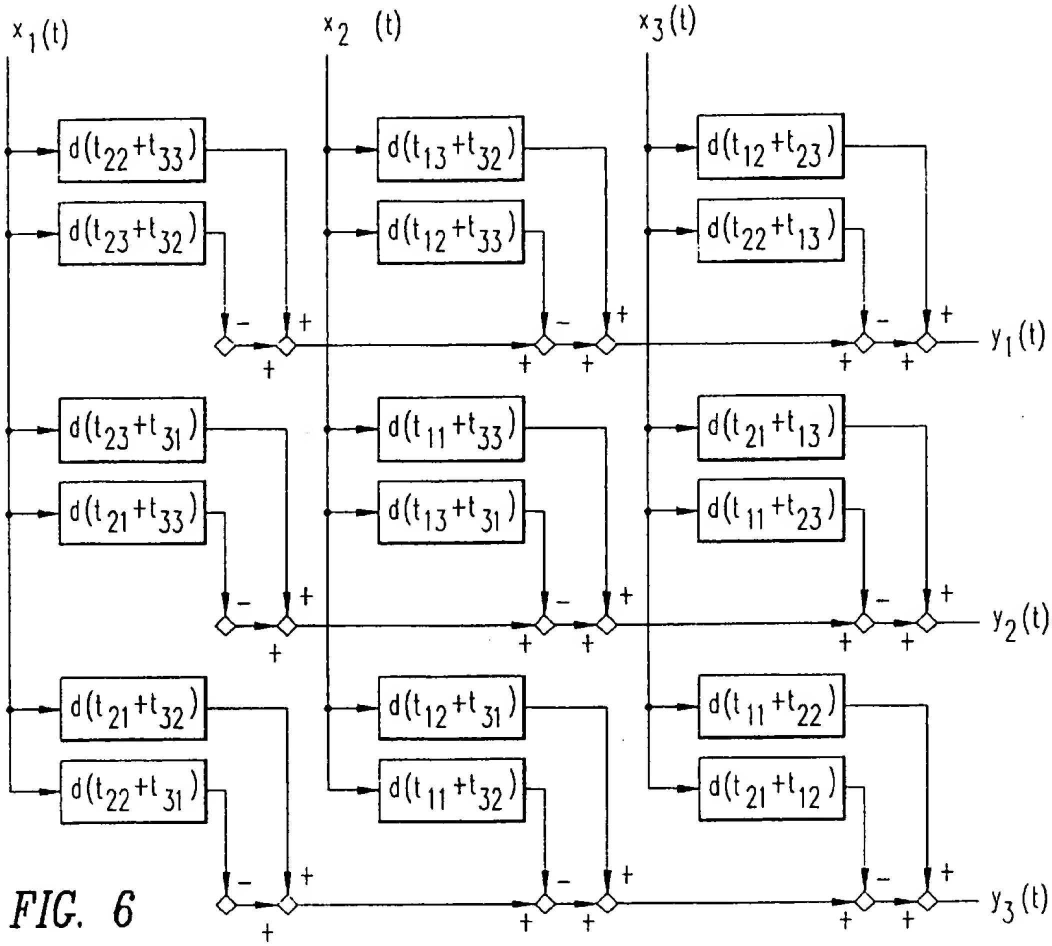

A+ echo$_1$ A
+ echo$_2$ B
B+ echo$_1$ B
+ echo$_2$ A
42
40
+
+
58
56
−
−
54
62
52
60
A+ echo$_1$ A
B+ echo$_1$ B

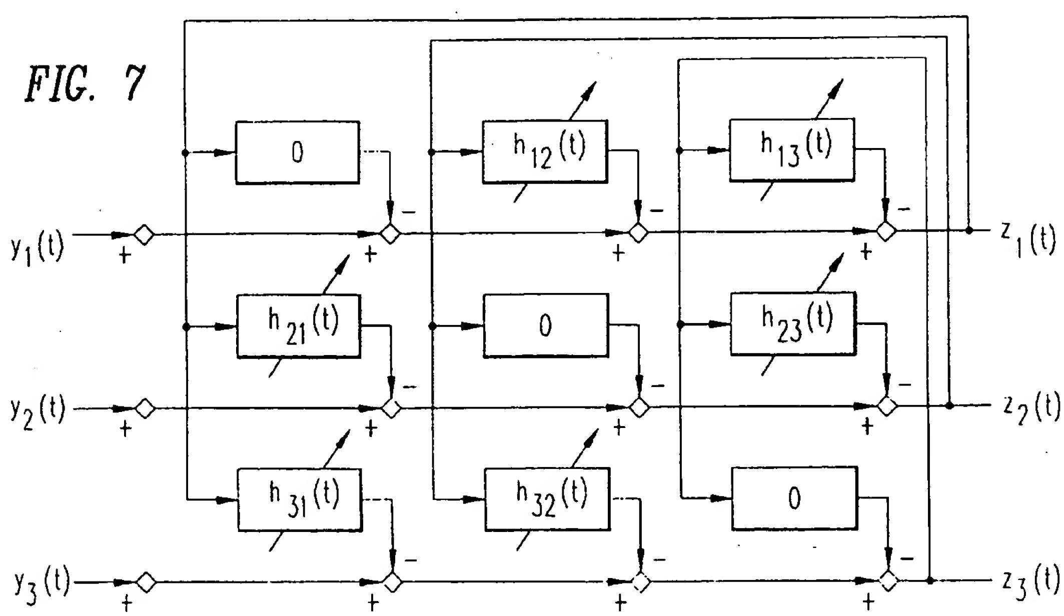
FIG. 5



x1(t)
x2 (t)
x3(t)
d(t22+t33)
d(t23+t32)
d(t13+t32)
d(t12+t33)
d(t12+t23)
d(t22+t13)
y1(t)
d(t23+t31)
d(t21+t33)
d(t11+t33)
d(t13+t31)
d(t21+t13)
d(t11+t23)
y2(t)
d(t21+t32)
d(t22+t31)
d(t12+t31)
d(t11+t32)
d(t11+t22)
d(t21+t12)
y3(t)

FIG. 6

FIG. 7



0
h12(t)
h13(t)
y1(t)
z1(t)
h21(t)
0
h23(t)
y2(t)
z2(t)
h31(t)
h32(t)
0
y3(t)
z3(t)

DIRECTIONAL ACOUSTIC SIGNAL PROCESSOR AND METHOD THEREFOR

TECHNICAL FIELD

This invention relates to the field of microphone-array signal processing, and more particularly to a two stage processor for extracting one substantially pure sound signal from a mixture of such signals even in the presence of echoes and reverberations.

BACKGROUND OF THE INVENTION

It is well known that a human being can focus attention on a single source of sound even in an environment that contains many such sources. This phenomenon is often called the "cocktail-party effect."

Considerable effort has been devoted in the prior art to solve the cocktail-party effect, both in physical devices and in computational simulations of such devices. One prior technique is to separate sound based on auditory scene analysis. In this analysis, vigorous use is made of assumptions regarding the nature of the sources present. It is assumed that a sound can be decomposed into small elements such as tones and bursts, which in turn can be grouped according to attributes such as harmonicity and continuity in time. Auditory scene analysis can be performed using information from a single microphone or from several microphones. For an early example of auditory scene analysis, see Weintraub (1984, 1985, 1986). Other prior art work related to sound separation by auditory scene analysis are due to Parsons (1976), von der Malsburg and Schneider (1986), Naylor and Porter (1991), and Mellinger (1991).

Techniques involving auditory scene analysis, although interesting from a scientific point of view as models of human auditory processing, are currently far too computationally demanding and specialized to be considered practical techniques for sound separation until fundamental progress is made.

Other techniques for separating sounds operate by exploiting the spatial separation of their sources. Devices based on this principle vary in complexity. The simplest such devices are microphones that have highly selective, but fixed patterns of sensitivity. A directional microphone, for example, is designed to have maximum sensitivity to sounds emanating from a particular direction, and can therefore be used to enhance one audio source relative to others (see Olson, 1967). Similarly, a close-talking microphone mounted near a speaker's mouth rejects distant sources (see, for example, the Knowles CF 2949 data sheet).

Microphone-array processing techniques related to separating sources by exploiting spatial separation of their sources are also well known and have been of interest for several decades. In one early class of microphone-array techniques, nonlinear processing is employed. In each output stream, some source direction of arrival, a "look direction," is assumed. The microphone signals are delayed to remove differences in time of flight from the look direction. Signals from any direction other than the look direction are thus misaligned in time. The signal in the output stream is formed, in essence, by "gating" sound fragments from the microphones. At any given instant, the output is chosen to be equal to one of the microphone signals. These techniques, exemplified by Kaiser and David (1960), by Mitchell et al. (1971), and by Lyon (1983), perform best when the undesired sources consist predominantly of impulse trains, as is the case with human speech. While these nonlinear techniques can be very computationally efficient and are of scientific interest as models of human cocktail-party processing, they do not have practical or commercial significance because of their inherent inability to bring about full suppression of unwanted sources. This inability originates from the incorrect assumption that at every instant in time, at least one microphone contains only the desired signal.

One widely known class of techniques in the prior art for linear microphone-array processing is often referred to as "classical beamforming" (Flanagan et al., 1985). As with the nonlinear techniques mentioned above, processing begins with the removal of time-of-flight differences among the microphone signals with respect to the look direction. In place of the "gating" scheme, the delayed microphone signals are simply averaged together. Thus, any signal from the look direction is represented in the output with its original power, whereas signals from other directions are relatively attenuated.

Classical beamforming was employed in a patented directional hearing aid invented by Widrow and Brearley (1988). The degree to which a classical beamformer is able to attenuate undesired sources relative to the desired source is limited by (1) the number of microphones in the array, and (2) the spatial extent of the array relative to the longest wavelength of interest present in the undesired sources. In particular, a classical beamformer cannot provide relative attenuation of frequency components whose wavelengths are larger than the array. For example, an array one foot wide cannot greatly attenuate frequency components below approximately 1 kHz.

Also known from the prior art is a class of active-cancellation algorithms, which is related to sound separation. However, it needs a "reference signal," i.e., a signal derived from only of one of the sources. For example, active noise-cancellation techniques (see data sheets for Bose® Aviation Headset, NCT proACTIVE® Series, and Sennheiser HDC451 Noiseguard® Mobile Headphone) reduce the contribution of noise to a mixture by filtering a known signal that contains only the noise, and subtracting it from the mixture. Similarly, echo-cancellation techniques such as those employed in full-duplex modems (Kelly and Logan, 1970; Gritton and Lin, 1984) improve the signal-to-noise ratio of an outgoing signal by removing noise due to echoes from the known incoming signal.

Techniques for active cancellation that do not require a reference signal are called "blind." They are now classified, based on the degree of realism of the underlying assumptions regarding the acoustic processes by which the unwanted signals reach the microphones. To understand the practical significance of this classification, recall a feature common to the principles by which active-cancellation techniques operate: the extent to which a given undesired source can be canceled by subtracting processed microphone signals depends ultimately on the exactness with which copies of the undesired source in the different microphones can be made to match one another. This depends, in turn, on how accurately the signal processing models the acoustic processes by which the unwanted signals reach the microphones.

One class of blind active-cancellation techniques may be called "gain-based": it is presumed that the waveform produced by each source is received by the microphones simultaneously, but with varying relative gains. (Directional microphones must be employed to produce the required differences in gain.) Thus, a gain-based system attempts to cancel copies of an undesired source in different microphone signals by applying relative gains to the microphone signals and subtracting, but never applying time delays or otherwise filtering. Numerous gain-based methods for blind active cancellation have been proposed; see Herault and Jutten (1986), Bhatti and Bibyk (1991), Cohen (1991), Tong et al. (1991), and Molgedey and Schuster (1994).

The assumption of simultaneity is violated when microphones are separated in space. A class of blind active-cancellation techniques that can cope with non-simultaneous mixtures may be called "delay-based": it is assumed that the waveform produced by each source is received by the various microphones with relative time delays, but without any other filtering. (See Morita, 1991 and Bar-Ness, 1993.) Under anechoic conditions, this assumption holds true for a microphone array that consists of omnidirectional microphones. However, this simple model of acoustic propagation from the sources to the microphones is violated when echoes and reverberation are present.

When the signals involved are narrowband, some gain-based techniques for blind active cancellation can be extended to employ complex gain coefficients (see Strube (1981), Cardoso (1989,1991), Lacoume and Ruiz (1992), Comon et al. (1994)) and can therefore accommodate, to a limited degree, time delays due to microphone separation as well as echoes and reverberation. These techniques can be adapted for use with audio signals, which are broadband, if the microphone signals are divided into narrowband components by means of a filter bank. The frequency bands produced by the filter bank can be processed independently, and the results summed (for example, see Strube (1981) or the patent of Comon (1994)). However, they are computationally intensive relative to the present invention because of the duplication of structures across frequency bands, are slow to adapt in changing situations, are prone to statistical error, and are extremely limited in their ability to accommodate echoes and reverberation.

The most realistic active-cancellation techniques currently known are "convolutive": the effect of acoustic propagation from each source to each microphone is modeled as a convolutive filter. These techniques are more realistic than gain-based and delay-based techniques because they explicitly accommodate the effects of inter-microphone separation, echoes and reverberation. They are also more general since, in principle, gains and delays are special cases of convolutive filtering.

Convolutive active-cancellation techniques have recently been described by Jutten et al. (1992), by Van Compernolle and Van Gerven (1992), by Platt and Faggin (1992), and by Dinc and Bar-Ness (1994). While these techniques have been used to separate mixtures constructed by simulation using oversimplified models of room acoustics, to the best of our knowledge none of them has been applied successfully to signals mixed in a real acoustic environment. The simulated mixtures used by Jutten et al., by Platt and Faggin, and by Dinc and Bar-Ness differ from those that would arise in a real room in two respects. First, the convolutive filters used in the simulations are much shorter than those appropriate for modeling room acoustics; they allowed for significant indirect propagation of sound over only one or two feet, compared with tens of feet typical of echoes and reverberation in an office. Second, the mixtures used in the simulations were partially separated to begin with, i.e., the crosstalk between the channels was weak. In practice, the microphone signals must be assumed to contain strong crosstalk unless the microphones are highly directional and the geometry of the sources is constrained.

To overcome some of the limitations of the convolutive active-cancellation techniques named above, the present invention employs a two-stage architecture. Its two-stage architecture is substantially different from other two-stage architectures found in prior art.

A two-stage signal processing architecture is employed in a Griffiths-Jim beamformer (Griffiths and Jim, 1982). The first stage of a Griffiths-Jim beamformer is delay-based: two microphone signals are delayed to remove time-of-flight differences with respect to a given look direction, and in contrast with classical beamforming, these delayed microphone signals are subtracted to create a reference noise signal. In a separate channel, the delayed microphone signals are added, as in classical beamforming, to create a signal in which the desired source is enhanced relative to the noise. Thus, the first stage of a Griffiths-Jim beamformer produces a reference noise signal and a signal that is predominantly desired source. The noise reference is then employed in the second stage, using standard active noise-cancellation techniques, to improve the signal-to-noise ratio in the output.

The Griffiths-Jim beamformer suffers from the flaw that under reverberant conditions, the delay-based first stage cannot construct a reference noise signal devoid of the desired signal, whereas the second stage relies on the purity of that noise reference. If the noise reference is sufficiently contaminated with the desired source, the second stage suppresses the desired source, not the noise (Van Compernolle, 1990). Thus, the Griffiths-Jim beamformer incorrectly suppresses the desired signal under conditions that are normally considered favorable: when the signal-to-noise ratio in the microphones is high.

Another two-stage architecture is described by Nájar et al. (1994). Its second stage employs blind convolutive active cancellation. However, its first stage differs significantly from the first stage of the Griffiths-Jim beamformer. It attempts to produce separated outputs by adaptively filtering each microphone signal in its own channel. When the sources are spectrally similar, filters that produce partially separated outputs after the first stage are unlikely to exist.

Thus, it is desirable to provide an architecture for separation of sources that avoids the difficulties exhibited by existing techniques.

SUMMARY OF THE INVENTION

An audio signal processing system for processing acoustic waves from a plurality of sources, comprising a plurality of spaced apart transducer means for receiving acoustic waves from the plurality of sources, including echoes and reverberations thereof. The transducer means generates a plurality of acoustic signals in response thereto. Each of the plurality of transducer means receives acoustic waves from the plurality of sources including echoes and reverberations thereof, and generates one of the plurality of acoustic signals. A first processing means receives the plurality of acoustic signals and generates a plurality of first processed acoustic signals in response thereto. In the absence of echoes and reverberations of the acoustic waves from the plurality of sources, each of the first processed acoustic signals represent acoustic waves from only one different source. A

second processing means receives the plurality of first processed acoustic signals and generates a plurality of second processed acoustic signals in response thereto. In the presence of echoes and reverberations of the acoustic waves from the plurality of sources, each of the second processed acoustic signals represent acoustic waves from only one different source.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a schematic block diagram of an embodiment of an acoustic signal processor of the present invention, using two microphones.

FIG. 2 is a schematic block diagram of an embodiment of the direct-signal separator portion, i.e., the first stage of the processor shown in FIG. 1.

FIG. 3 is a schematic block diagram of an embodiment of the crosstalk remover portion, i.e., the second stage of the processor shown in FIG. 1.

FIG. 4 is an overview of the delay in the acoustic waves arriving at the direct signal separator portion of the signal processor of FIG. 1, and showing the separation of the signals.

FIG. 5 is an overview of a portion of the crosstalk remover of the signal processor of FIG. 1 showing the removal of the crosstalk from one of the signal channels.

FIG. 6 is a detailed schematic block diagram of an embodiment of a direct-signal separator using three microphones.

FIG. 7 is a detailed schematic block diagram of an embodiment of a crosstalk remover suitable using three microphones.

DETAILED DESCRIPTION OF THE INVENTION

The present invention is a device that mimics the cocktail-party effect using a plurality of microphones with as many output audio channels, and a signal-processing module. When situated in a complicated acoustic environment that contains multiple audio sources with arbitrary spectral characteristics, it supplies output audio signals, each of which contains sound from at most one of the original sources. These separated audio signals can be used in a variety of applications, such as hearing aids or voice-activated devices.

FIG. 1 is a schematic diagram of a signal separator processor of one embodiment of the present invention. As previously discussed, the signal separator processor of the present invention can be used with any number of microphones. In the embodiment shown in FIG. 1, the signal separator processor receives signals from a first microphone **10** and a second microphone **12**, spaced apart by about two centimeters. As used herein, the microphones **10** and **12** include transducers (not shown), their associated preamplifiers (not shown), and A/D converters **22** and **24** (shown in FIG. 2).

The microphones **10** and **12** in the preferred embodiment are omnidirectional microphones, each of which is capable of receiving acoustic wave signals from the environment and for generating a first and a second acoustic electrical signal **14** and **16** respectively. The microphones **10** and **12** are either selected or calibrated to have matching sensitivity. The use of matched omnidirectional microphones **10** and **12**, instead of directional or other microphones leads to simplicity in the direct-signal separator **30**, described below. In the preferred embodiment, two Knowles EM-3046 omnidirectional microphones were used, with a separation of 2 centimeters. The pair was mounted at least 25 centimeters from any large surface in order to preserve the omnidirectional nature of the microphones. Matching was achieved by connecting the two microphone outputs to a stereo microphone preamplifier and adjusting the individual channel gains so that the preamplifier outputs were closely matched. The preamplifier outputs were each digitally sampled at 22,050 samples per second, simultaneously. These sampled electrical signals **14** and **16** are supplied to the direct signal separator **30** and to a Direction of Arrival (DOA) estimator **20**.

The direct-signal separator **30** employs information from a DOA estimator **20**, which derives its estimate from the microphone signals. In a different embodiment of the invention, DOA information could come from an source other than the microphone signals, such as direct input from a user via an input device.

The direct signal separator **30** generates a plurality of output signals **40** and **42**. The direct signal separator **30** generates as many output signals **40** and **42** as there are microphones **10** and **12**, generating as many input signals **14** and **16** as are supplied to the direct signal separator **30**. Assuming that there are two sources, A and B, generating acoustic wave signals in the environment in which the signal processor **8** is located, then each of the microphones **10** and **12** would detect acoustic waves from both sources. Hence, each of the electrical signals **14** and **16**, generated by the microphones **10** and **12**, respectively, contains components of sound from sources A and B.

The direct-signal separator **30** processes the signals **14** and **16** to generate the signals **40** and **42** respectively, such that in anechoic conditions (i.e., the absence of echoes and reverberations), each of the signals **40** and **42** would be of an electrical signal representation of sound from only one source. In the absence of echoes and reverberations, the electrical signal **40** would be of sound only from source A, with electrical signal **42** being of sound only from source B, or vice versa. Thus, under anechoic conditions the direct-signal separator **30** can bring about full separation of the sounds represented in signals **14** and **16**. However, when echoes and reverberation are present, the separation is only partial.

The output signals **40** and **42** of the direct signal separator **30** are supplied to the crosstalk remover **50**. The crosstalk remover **50** removes the crosstalk between the signals **40** and **42** to bring about fully separated signals **60** and **62** respectively. Thus, the direct-signal separator **30** and the crosstalk remover **50** play complementary roles in the system **8**. The direct-signal separator **30** is able to bring about full separation of signals mixed in the absence of echoes and reverberation, but produces only partial separation when echoes and reverberation are present. The crosstalk remover **50** when used alone is often able to bring about full separation of sources mixed in the presence of echoes and reverberation, but is most effective when given inputs **40** and **42** that are partially separated.

After some adaptation time, each output **60** and **62** of the crosstalk remover **50** contains the signal from only one sound source: A or B. Optionally, these outputs **60** and **62** can be connected individually to post filters **70** and **72**, respectively, to remove known frequency coloration produced by the direct signal separator **30** or the crosstalk remover **50**. Practitioners skilled in the art will recognize that there are many ways to remove this known frequency coloration; these vary in terms of their cost and effective-

ness. An inexpensive post filtering method, for example, consists of reducing the treble and boosting the base.

The filters **70** and **72** generate output signals **80** and **82**, respectively, which can be used in a variety of applications. For example, they may be connected to a switch box and then to a hearing aid.

Referring to FIG. 2 there is shown one embodiment of the direct signal separator **30** portion of the signal processor **8** of the present invention. The microphone transducers generate input signals **11** and **13**, which are sampled and digitized, by clocked sample-and-hold circuits followed by analog-to-digital converters **22** and **24**, respectively, to produce sampled digital signals **14** and **16** respectively.

The digital signal **14** is supplied to a first delay line **32**. In the preferred embodiment, the delay line **32** delays the digitally sampled signal **14** by a non-integral multiple of the sampling interval T, which was 45.35 microseconds given the sampling rate of 22,050 samples per second. The integral portion of the delay was implemented using a digital delay line, while the remaining subsample delay of less than one sample interval was implemented using a non-causal, truncated sinc filter with 41 coefficients. Specifically, to implement a subsample delay of t, given that t<T, the following filter is used:

$$y(n)=\sum_{k=-20}^{20} w(k)\ x(n-k)$$

where x(n) is the signal to be delayed, y(n) is the delayed signal, and w(k) {k=-20, -19, . . . 19,20} are the 41 filter coefficients. The filter coefficients are determined from the subsample delay t as follows:

$$w(k)=(1/S)\ \mathrm{sinc}\{\pi[(t/T)-k]\}$$

where

$$\mathrm{sinc}(a)=\sin(a)/a \quad \text{if a not equal to 0}$$
$$=1 \qquad \text{otherwise,}$$

and S is a normalization factor given by

$$S=\sum_{k=-20}^{20} \mathrm{sinc}\{\pi[(t/T)-k]\}.$$

The output of the first delay line **32** is supplied to the negative input of a second combiner **38**. The first digital signal **14** is also supplied to the positive input of a first combiner **36**. Similarly, for the other channel, the second digital signal **16** is supplied to a second delay line **34**, which generates a signal which is supplied to the negative input of the first combiner **36**.

In the preferred embodiment, the sample-and-hold and A/D operations were implemented by the audio input circuits of a Silicon Graphics Indy workstation, and the delay lines and combiners were implemented in software running on the same machine.

However, other delay lines such as analog delay lines, surface acoustic wave delays, digital low-pass filters, or digital delay lines with higher sampling rates, may be used in place of the digital delay line **32**, and **34**. Similarly, other combiners, such as analog voltage subtractors using operational amplifiers, or special purpose digital hardware, may be used in place of the combiners **36** and **38**.

Schematically, the function of the direct signal separator **30** may be seen by referring to FIG. **4**. Assuming that there are no echoes or reverberations, the acoustic wave signal received by the microphone **12** is the sum of source B and a delayed copy of source A. (For clarity in presentation here and in the forthcoming theory section, time relationship between the sources A and B and the microphones **10** and **12** are described as if the electrical signal **14** generated by the microphone **10** were simultaneous with source A and the electrical signal **16** generated by the microphone **12** were simultaneous with source B. This determines the two-arbitrary additive time constants that one is free to choose in each channel.) Thus, the electrical signal **16** generated by the microphone **12** is an electrical representation of the sound source B plus a delayed copy of source A. Similarly, the electrical signal **14** generated by the microphone **10** is an electrical representation of the sound source A and a delayed copy of sound source B. By delaying the electrical signal **14** an appropriate amount, the electrical signal supplied to the negative input of the combiner **38** would represent a delayed copy of source A plus a further delayed copy of source B. The subtraction of the signal from the delay line **32** and digital signal **16** would remove the signal component representing the delayed copy of sound source A, leaving only the pure sound B (along with the further delayed copy of B).

The amount of delay to be set for each of the digital delay lines **32** and **34** can be supplied from the DOA estimator **20**. Numerous methods for estimating the relative time delays have been described in the prior art (for example, Schmidt, 1981; Roy et al., 1988; Morita, 1991; Allen, 1991). Thus, the DOA estimator **20** is well known in the art.

In a different embodiment, omni-directional microphones **10** and **12** could be replaced by directional microphones placed very close together. Then all delays would be replaced by multipliers; in particular, digital delay lines **32** and **34** would be replaced by multipliers. Each multiplier would receive the signal from its respective A/D converter and generate a scaled signal, which can be either positive or negative, in response.

A preferred embodiment of the crosstalk remover **50** is shown in greater detail in FIG. **3**. The crosstalk remover **50** comprises a third combiner **56** for receiving the first output signal **40** from the direct signal separator **30**. The third combiner **56** also receives, at its negative input, the output of a second adaptive filter **54**. The output of the third combiner **56** is supplied to a first adaptive filter **52**. The output of the first adaptive filter **52** is supplied to the negative input of the fourth combiner **58**, to which the second output signal **42** from the direct signal separator **30** is also supplied. The outputs of the third and fourth combiners **56** and **58** respectively, are the output signals **60** and **62**, respectively of the crosstalk remover **50**.

Schematically, the function of the crosstalk remover **50** may be seen by referring to FIG. **5**. The inputs **40** and **42** to the crosstalk remover **50** are the outputs of the direct-signal separator **30**. Let us assume that the direct-signal separator **30** has become fully adapted, i.e., (a) that the electrical signal **40** represents the acoustic wave signals of source B and its echoes and reverberation, plus echoes and reverberation of source A, and similarly (b) that the electrical signal **42** represents the acoustic wave signals of source A and its echoes and reverberation, plus echoes and reverberation of source B. Because the crosstalk remover **50** is a feedback network, it is easiest to analyze subject to the assumption that adaptive filters **52** and **54** are fully adapted, so that the electrical signals **62** and **60** already correspond to colored versions of B and A, respectively. The processing of the

electrical signal **60** by the adaptive filter **52** will generate an electrical signal equal to the echoes and reverberation of source B present in the electrical signal **42**; hence subtraction of the output of adaptive filter **52** from the electrical signal **42** leaves output signal **62** with signal components only from source A. Similarly, the processing of the electrical signal **62** by the adaptive filter **54** will generate an electrical signal equal to the echoes and reverberation of source A present in the electrical signal **40**; hence subtraction of the output of adaptive filter **54** from the electrical signal **40** leaves output signal **60** with signal components only from source B.

Theory

It is assumed, solely for the purpose of designing the direct-signal separator **30**, that the microphones **10** and **12** are omnidirectional and matched in sensitivity.

Under anechoic conditions, the signals $x_1(t)$ and $x_2(t)$, which correspond to the input signals, received by microphones **10** and **12**, respectively, may be modeled as

$$x_1(t)=w_1(t)+w_2(t-\tau_2)$$

$$x_2(t)=w_2(t)+w_1(t-\tau_1),$$

where $w_1(t)$ and $w_2(t)$ are the original source signals, as they reach microphones **10** and **12**, respectively, and τ_1 and τ_2 are unknown relative time delays, each of which may be positive or negative.

Practitioners experienced in the art will recognize that bounded "negative" time delays can be achieved by adding a net time delay to the entire system.

The relative time delays τ_1 and τ_2 are used to form outputs $y_1(t)$ and $y_2(t)$, which correspond to signals **40** and **42**:

$$y_1(t)=x_1(t)-x_2(t-\tau_2)=w_1(t)-w_1(t-(\tau_1+\tau_2))$$

$$y_2(t)=x_2(t)-x_1(t-\tau_1)=w_2(t)-w_2(t-(\tau_1+\tau_2))$$

As depicted in FIG. **2**, these operations are accomplished by time-delay units **32** and **34**, and combiners **36** and **38**.

Under anechoic conditions, these outputs **40** and **42**, would be fully separated; i.e., each output **40** or **42** would contain contributions from one source alone. However under echoic conditions these outputs **40** and **42** are not fully separated.

Under echoic and reverberant conditions, the microphone signals $x_1(t)$ and $x_2(t)$, which correspond to input signals received by the microphones **10** and **12**, respectively, may be modeled as

$$x_1(t)=w_1(t)+w_2(t-\tau_2)+k_{11}'(t)*w_1(t)+k_{12}'(t)*w_2(t)$$

$$x_2(t)=w_2(t)+w_1(t-\tau_1)+k_{21}'(t)*w_1(t)+k_{22}'(t)*w_2(t),$$

where the symbol "*" denotes the operation of convolution, and the impulse responses $k_{11}'(t)$, $k_{12}'(t)$, $k_{21}'(t)$, and $k_{22}'(t)$ incorporate the effects of echoes and reverberation.

Specifically, $k_{11}'(t)*w_1(t)$ represents the echoes and reverberations of source **1** ($w_1(t)$) as received at input **1** (microphone **10**), $k_{12}'(t)*w_2(t)$ represents the echoes and reverberations of source **2** ($w_2(t)$) as received at input **1** (microphone **10**), $k_{21}'(t)*w_1(t)$ represents the echoes and reverberations of source **1** ($w_1(t)$) as received at input **2** (microphone **12**), and $k_{22}'(t)*w_2(t)$ represents the echoes and reverberations of source **2** ($w_2(t)$) as received at input **2** (microphone **12**).

In consequence of the presence of echoes and reverberation, the outputs **40** and **42** from the direct-signal separator **30** are not fully separated, but instead take the form

$$y_1(t)=x_1(t)-x_2(t-\tau_2)=w_1(t)-w_1(t-(\tau_1+\tau_2))+k_{11}(t)*w_1(t)+k_{12}(t)*w_2(t)$$

$$y_2(t)=x_2(t)-x_1(t-\tau_1)=w_2(t)-w_2(t-(\tau_1+\tau_2))+k_{21}(t)*w_1(t)+k_{22}(t)*w_2(t)$$

where the filters $k_{11}(t)$, $k_{12}(t)$, $k_{21}(t)$, and $k_{22}(t)$ are related to $k_{11}'(t)$, $k_{12}'(t)$, $k_{21}'(t)$, and $k_{22}'(t)$ by time shifts and linear combinations. Specifically,

$$k_{11}(t)=k_{11}'(t)-k_{21}'(t-\tau_2),$$

$$k_{12}(t)=k_{12}'(t)-k_{22}'(t-\tau_2),$$

$$k_{21}(t)=k_{21}'(t)-k_{11}'(t-\tau_1),\text{ and}$$

$$k_{22}(t)=k_{22}'(t)-k_{12}'(t-\tau_1).$$

Note that $y_1(t)$ is contaminated by the term $k_{12}(t)*w_2(t)$, and that $y_2(t)$ is contaminated by the term $k_{21}(t)*w_1(t)$.

Several possible forms of the crosstalk remover have been described as part of the background of this invention, under the heading of convolutive blind source separation. In the present embodiment, the crosstalk remover forms discrete time sampled outputs **60** and **62** thus:

$$z_1(n)=y_1(n)-\sum_{k=1}^{1000}h_2(k)z_2(n-k)$$

$$z_2(n)=y_2(n)-\sum_{k=1}^{1000}h_1(k)z_1(n-k)$$

where the discrete time filters h_1 and h_2 correspond to elements **52** and **54** in FIG. **3** and are estimated adaptively. The filters h_1 and h_2 are strictly causal, i.e., they operate only on past samples of z_1 and Z_2. This structure was described independently by Jutten et al. (1992) and by Platt and Faggin (1992).

The adaptation rule used for the filter coefficients in the preferred embodiment is a variant of the LMS rule ("Adaptive Signal Processing," Bernard Widrow and Samuel D. Stearns, Prentice-Hall, Englewood Cliffs, N.J., 1985, p 99). The filter coefficients are updated at every time-step n, after the new values of the outputs $z_1(n)$ and $Z_2(n)$ have been calculated. Specifically, using these new values of the outputs, the filter coefficients are updated as follows :

$$h_1(k)[\text{new}]=h_1(k)[\text{old}]+m\,z_2(n)z_1(n-k)\quad k=1,2,\ldots,1000$$

$$h_2(k)[\text{new}]=h_2(k)[\text{old}]+m\,z_1(n)z_2(n-k)\quad k=1,2,\ldots,1000$$

where m is a constant that determines the rate of adaptation of the filter coefficients, e.g. 0.15 if the input signals **10** and **12** were normalized to lie in the range $-1\leqq x(t)\leqq +1$. One skilled in the art will recognize that the filters $_1$ and h_2 can be implemented in a variety of ways, including FIRs and lattice IIRs.

As described, the direct-signal separator **30** and crosstalk remover **50** adaptively bring about full separation of two sound sources mixed in an echoic, reverberant acoustic environment. However, the output signals $z_1(t)$ and $z_2(t)$ may be unsatisfactory in practical applications because they are colored versions of the original sources $w_1(t)$ and $w_2(t)$ i.e.,

$$z_1=\zeta_1(t)*w_1(t)$$

$$z_2=\zeta_2(t)*w_2(t)$$

where $\zeta_1(t)$ and $\zeta_2(t)$ represent the combined effects of the echoes and reverberations and of the various known signal

transformations performed by the direct-signal separator **30** and crosstalk remover **50**.

As an optional cosmetic improvement for certain commercial applications, it may be desirable to append filters **70** and **72** to the network. The purpose of these filters is to undo the effects of filters $\zeta_1(t)$ and $\zeta_2(t)$. As those familiar with the art will realize, a large body of techniques for performing this inversion to varying and predictable degrees of accuracy currently exist.

The embodiment of the signal processor **8** has been described in FIGS. **1–5** as being useful with two microphones **10** and **12** for separating two sound sources, A and B. Clearly, the invention is not so limited. The forthcoming section describes how more than two microphones and sound sources can be accomodated.

General Case with M Microphones and M Sources

The invention is able to separate an arbitrary number M of simultaneous sources, as long as they are statistically independent, if there are at least M microphones.

Let $w_j(t)$ be the j'th source signal and $x_i(t)$ be the i'th microphone (mic) signal. Let t_{ij} be the time required for sound to propagate from source j to mic i, and let $d(t_{ij})$ be the impulse response of a filter that delays a signal by t_{ij}. Mathematically, $d(t_{ij})$ is the unit impulse delayed by t_{ij}, that is

$$d(t_{ij})=\delta(t-t_{ij})$$

where $\delta(t)$ is the unit impulse function ("Circuits, Signals and Systems", by William McC. Siebert. The MIT Press, McGraw Hill Book Company, 1986, p. 319).

In the absence of echoes and reverberation, the i'th mic signal $x_i(t)$ can be expressed as a sum of the appropriately delayed j source signals

$$x_i(t)=\sum_{j=1}^{M} d(t_{ij})*w_j(t)$$

Matrix representation allows a compact representation of this equation for all M mic signals:

$$X(t)=D(t)*W(t)$$

where X(t) is an M-element column vector whose i'th element is the i'th mic signal $x_i(t)$, D(t) is an M×M element square matrix whose ij'th element (ie., the element in the i'th row and j'th column) is $d(t_{ij})$, and W(t) is an M-element column vector whose j'th element is the j'th source signal $w_j(t)$. Specifically,

$$X(t)=\begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_M(t) \end{bmatrix};$$

$$D(t)=\begin{bmatrix} d(t_{11}) & d(t_{12}) & \cdots & d(t_{1M}) \\ d(t_{21}) & d(t_{22}) & \cdots & d(t_{2M}) \\ \cdots & \cdots & & \cdots \\ \cdots & \cdots & & \cdots \\ d(t_{MI}) & d(t_{M2}) & \cdots & d(t_{MM}) \end{bmatrix};$$

-continued

$$W(t)=\begin{bmatrix} w_1(t) \\ w_2(t) \\ \vdots \\ w_M(t) \end{bmatrix}$$

For each source $w_j(t)$, if the delays t_{ij} for i=1,2, . . . , M to the M mics are known (up to an arbitrary constant additive factor that can be different for each source), then M signals $y_j(t)$, j=1,2, . . . , M, that each contain energy from a single but different source $w_j(t)$, can be constructed from the mic signals $x_i(t)$ as follows:

$$y(t)=adjD(t)*x(t),$$

where

$$Y(t)=\begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_M(t) \end{bmatrix}$$

is the M-element column vector whose j'th element is the separated signal $y_j(t)$, and adjD(t) is the adjugate matrix of the matrix D(t). The adjugate matrix of a square matrix is the matrix obtained by replacing each element of the original matrix by its cofactor, and then transposing the result ("Linear Systems", by Thomas Kailath, Prentice Hall, Inc., 1980, p. 649). The product of the adjugate matrix and the original matrix is a diagonal matrix, with each element along the diagonal being equal to the determinant of the original matrix. Thus,

$$\begin{aligned} Y(t) &= adjD(t)*X(t) \\ &= adjD(t)*D(t)*W(t) \\ &= \begin{bmatrix} |D(t)| & 0 & 0\cdots 0 \\ 0 & |D(t)| & 0\cdots 0 \\ \vdots & \ddots & \vdots \\ 0 & 0 & \cdots |D(t)| \end{bmatrix} * W(t) \\ &= |D(t)|*W(t) \end{aligned}$$

where $|D(t)|$ is the determinant of D(t). Thus,

$$y_j(t)=|D(t)*w_j(t) \text{ for } j=1,2, \ldots ,M$$

$y_j(t)$ is a "colored" or filtered version of $w_j(t)$ because of the convolution by the filter impulse response $|D(t)|$. If desired, this coloration can be undone by post filtering the outputs by a filter that is the inverse of $|D(t)|$. Under certain circumstances, determined by the highest frequency of interest in the source signals and the separation between the mics, the filter $|D(t)|$ may have zeroes at certain frequencies; these make it impossible to exactly realize the inverse of the filter $|D(t)|$. Under these circumstances any one of the numerous techniques available for approximating filter inverses (see, for example, "Digital Filters and Signal Processing", by Leland B. Jackson, Kluwer Academic Publishers, 1986, p.146) may be used to derive an approximate filter with which to do the post filtering.

The delays t_{ij} can be estimated from the statistical properties of the mic signals, up to a constant additive factor that can be different for each source. This is the subject of a

co-pending patent application by the same inventors, filed even date herewith. Alternatively, if the position of each microphone and each source is known, then the delays t_{ij} can be calculated exactly. For any source that is distant, i.e., many times farther than the greatest separation between the mics, only the direction of the source is needed to calculate its delays to each mic, up to an arbitrary additive constant.

The first stage of the processor **8**, namely the direct signal separator **30**, uses the estimated delays to construct the adjugate matrix adjD(t), which it applies to the microphone signals X(t) to obtain the outputs Y(t) of the first stage, given by:

$$Y(t)=adjD(t)*X(t).$$

In the absence of echoes and reverberations, each output $y_j(t)$ contains energy from a single but different source $w_j(t)$.

When echoes and reverberation are present, each mic receives the direct signals from the sources as well as echoes and reverberations from each source. Thus

$$x_i(t)=\sum_{j=1}^{M} d(t_{ij})*w_j(t)+\sum_{j=1}^{M} e_{ij}(t)*w_j(t) \quad \text{for } i=1,2,\ldots,M$$

where $e_{ij}(t)$ is the impulse response of the echo and reverberation path from the j'th source to the i'th mic. All M of these equations can be represented in compact matrix notation by

$$X(t)=D(t)*W(t)+E(t)*W(t)$$

where E(t) is the M×M matrix whose ijth element is the fitter $e_{ij}(t)$.

If the mic signals are now convolved with the adjugate matrix of D(t), instead of obtaining separated signals we obtain partially separated signals:

$$\begin{aligned} Y(t) &= adjD(t)*X(t) \\ &= |D(t)|*W(t)+adjD(t)*E(t)*W(t) \end{aligned}$$

Notice that each $y_j(t)$ contains a colored direct signal from a single source, as in the case with no echoes, and differently colored components from the echoes and reverberations of every source, including the direct one.

The echoes and reverberations of the other sources are removed by the second stage of the network, namely the crosstalk remover **50**, which generates each output as follows:

$$z_j(t)=y_j(t)-\sum_{\substack{k=1 \\ k\neq j}}^{M} h_{jk}(t)*z_k(t) \quad \text{for } j=1,2,\ldots,M$$

where the entities $h_{jk}(t)$ are causal adaptive filters. (The term "causal" means that $h_{jk}(t)=0$ for $t\leqq 0$.) In matrix form these equations are written as

$$Z(t)=Y(t)-H(t)*Z(t)$$

where Z(t) is the M-element column vector whose j'th element is $z_j(t)$, and H(t) is an M×M element matrix whose diagonal elements are zero and whose off diagonal elements are the causal, adaptive filters $h_{jk}(t)$.

These filters are adapted according to a rule that is similar to the Least Mean Square update rule of adaptive filter theory ("Adaptive Signal Processing," Bernard Widrow and Samuel D. Stearns, Prentice-Hall, Englewood Cliffs, N.J., 1985, p. 99).

This is most easily illustrated in the case of a discrete time system.

Illustrative Weight Update Methodology for Use with a Discrete Time Representation

First, we replace the time parameter by a discrete time index n. Second, we use the notation H(n)[new] to indicate the value of H(n) in effect just before computing new outputs at time n. At each time step n, the outputs Z(n) are computed according to

$$Z(n)=Y(n)-H(n)[\text{new}]*Z(n)$$

Note that the convolution on the right hand side involves only past values of Z, ie Z(n−1), Z(n−2), . . . , Z(n−N), because the filters that are the elements of H are causal. (N is defined to be the order of the filters in H).

Now new values are computed for the coefficients of the filters that are the elements of H. These will be used at the next time step. Specifically, for each j and each k, with j≠k, perform the following:

$$\begin{aligned} h_{jk}(u)[\text{old}] &= h_{jk}(u)[\text{new}] & u=1,2,\ldots,M \\ h_{jk}(u)[\text{new}] &= h_{jk}(u)[\text{old}]+\mu_{jk}z_j(n)z_k(n-u) & u=1,2,\ldots,M \end{aligned}$$

The easiest way to understand the operation of the second stage is to observe that the off-diagonal elements of H(t) have zero net change per unit time when the products like $z_j(t)z_k(t-u)$ are zero on average. Because the sources in W are taken to be statistically independent of each other, those products are zero on average when each output $z_j(t)$ has become a colored version of a different source, say $w_j(t)$. (The correspondence between sources and outputs might be permuted so that the numbering of the sources does not match the numbering of the outputs.)

More specifically, let $Z(t)=\psi(t)*W(t)$. From the preceding paragraph, equilibrium is achieved when $\psi(t)$ is diagonal. In addition, it is required that:

$$\begin{aligned} Z(t) &= Y(t)-H(t)*\Psi(t)*W(t) \\ &= |D(t)|*W(t)+adjD(t)*E(t)*W(t)-H(t)*\Psi(t)*W(t) \\ &= (|D(t)|I+adjD(t)*E(t)*-H(t)*\Psi(t))*W(t) \end{aligned}$$

so that

$$\Psi(t)=|D(t)|I+adjD(t)*E(t)-H(t)*\Psi(t)$$

$$\Psi(t)=[1+H(t)]^{-1}[|D(t)|I+adjD(t)*E(t)]$$

This relation determines the coloration produced by the two stages of the system, taken together.

An optional third stage can use any one of numerous techniques available to reduce the amount of coloration on any individual output.

Example of General Case with M=3, i.e. with 3 Mics and 3 Sources

In the case where there are 3 mics and 3 sources, the general matrix equation

$$X=D*W$$

becomes

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} d(t_{11}) & d(t_{12}) & d(t_{13}) \\ d(t_{21}) & d(t_{22}) & d(t_{23}) \\ d(t_{31}) & d(t_{32}) & d(t_{33}) \end{bmatrix} * \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}$$

If the delays t_{ij} are known, then the adjugate matrix of D(t) is given by

$$adjD = \begin{bmatrix} d(t_{22}+t_{33})-d(t_{23}+t_{32}) & d(t_{13}+t_{32})-d(t_{12}+t_{33}) & d(t_{12}+t_{23})-d(t_{22}+t_{13}) \\ d(t_{23}+t_{31})-d(t_{21}+t_{33}) & d(t_{11}+t_{33})-d(t_{13}+t_{31}) & d(t_{21}+t_{13})-d(t_{11}+t_{23}) \\ d(t_{21}+t_{32})-d(t_{22}+t_{31}) & d(t_{12}+t_{31})-d(t_{11}+t_{32}) & d(t_{11}+t_{22})-d(t_{21}+t_{12}) \end{bmatrix}$$

Note that adding a constant delay to the delays associated with any column of D(t) leaves the adjugate matrix unchanged. This is why the delays from a source to the three mics need only be estimated up to an arbitrary additive constant.

The output of the first stage, namely the direct signal separator **30**, is formed by convolving the mic signals with the adjugate matrix.

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = adjD * \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

The network that accomplishes this is shown in FIG. **6**.

In the absence of echoes and reverberations, the outputs of the first stage are the individual sources, each colored by the determinant of the delay matrix.

$$\begin{bmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{bmatrix} = adjD * D(t)$$

$$\begin{bmatrix} w_1(t) \\ w_2(t) \\ w_3(t) \end{bmatrix} = |D(t)| * \begin{bmatrix} w_1(t) \\ w_2(t) \\ w_3(t) \end{bmatrix}$$

In the general case when echoes and reverberation are present, each output of the first stage also contains echoes and reverberations from each source. The second stage, namely the cross talk remover **50**, consisting of a feedback network of adaptive filters, removes the effects of these unwanted echoes and reverberations to produce outputs that each contain energy only one different source, respectively. The matrix equation of the second stage

$$Z=Y-H*Z$$

becomes

$$\begin{bmatrix} z_1(t) \\ z_2(t) \\ z_3(t) \end{bmatrix} = \begin{bmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{bmatrix} - \begin{bmatrix} 0 & h_{12}(t) & h_{13}(t) \\ h_{21}(t) & 0 & h_{23}(t) \\ h_{31}(t) & h_{32}(t) & 0 \end{bmatrix} * \begin{bmatrix} z_1(t) \\ z_2(t) \\ z_3(t) \end{bmatrix}$$

where each h_{ij} is a causal adaptive filter. The network that accomplishes this is shown in FIG. **7**.

Conclusion

It should be noted that the number of microphones and associated channels of signal processing need not be as large as the total number of sources present, as long as the number of sources emitting a significant amount of sound at any given instant in time does not exceed the number of microphones. For example, if during one interval of time only sources A and B emit sound, and in another interval of time only sources B and C emit sound, then during the first interval the output channels will correspond to A and B respectively, and during the second interval the output channels will correspond to B and C respectively.

As previously discussed, in the preferred embodiment of the present invention, the invention has been implemented in software, as set forth in the microfiche appendix. The software code is written in the C++ language for execution on a workstation from Silicon Graphics. However, as previously discussed, hardware implementation of the present invention is also contemplated to be within the scope of the present invention. Thus, for example, the direct signal separator **30** and the crosstalk remover **50** can be a part of a digital signal processor, or can be a part of a general purpose computer, or can be a part of analog signal processing circuitry. In addition, the present invention is not limited to the processing of acoustic waves. It can be used to process signals, having a delay with the problem of separation of the signals from the sources.

There are many advantages and differences of the present invention from the prior art.

1. Although the present invention is similar in objective to sound separation based on auditory scene analysis, it differs from them in principle and in technique.
2. In contrast with approaches based on auditory scene analysis, this invention separates sounds by exploiting the separation of their sources in space and their statistical independence.
3. The present invention differs from directional microphones in that few presuppositions are made with regard to the locations of sound sources relative to the device. The present device need not be pointed at or mounted close to a source of interest. The necessary selectivity is brought about by processing of signals captured by a microphone array, i.e., a collection of microphones. Moreover, the selectivity attained is much greater: a directional microphone cannot completely suppress any source of sound.
4. While the prior art nonlinear techniques can be very computationally efficient and are of scientific interest as models of human cocktail-party processing, they are of less practical or commercial significance than the present invention because of their inherent inability to bring about full suppression of unwanted sources. This inability originates from the incorrect assumption that at every instant in time, at least one microphone contains only the desired signal. The present invention differs from these nonlinear techniques in that linear operations (addition, subtraction, and filtering) are employed to cancel unwanted sources.
5. The present invention is an example of "active cancellation." In contrast with classical beamforming, which aligns copies of a desired source in time and adds them together, active cancellation matches copies of undesired sources and subtracts them to bring about cancellation. ("Matching" the copies of the undesired sources generally

entails more than simply re-aligning them in time; usually, it involves re-shaping them by filtering.) Bearing this simplified explanation in mind, it may be seen that the degree of selectivity achieved by active cancellation is determined by factors rather different from those important in classical beamforming. In active cancellation, the degree of selectivity is determined by the exactness with which identical copies of unwanted sources can be created from different microphones for subsequent cancellation. In contrast with a classical beamformer, a sound-separation device that employs active cancellation can in principle remove one undesired source completely using just two microphones.

6. The present invention also does not need a reference sound.
7. In contrast with active-cancellation algorithms that require a reference signal, the present invention operates "blindly": it accommodates the more difficult case in which all of the signals directly available for processing, i.e., the microphone signals, are assumed to be mixtures of sources. The present invention is a method for bringing about blind separation of sources, given the direction of arrival (DOA) of each source. This DOA information may be obtained in a variety of ways, for example by direct specification from a human user or by statistical estimation from the microphone signals. Precisely how DOA information is obtained is immaterial in the context of the present invention; what is important is how DOA information is used to bring about source separation.
8. The present invention differs from gain-based active cancellation in that it requires no assumption of simultaneity of signal reception at all of the microphones.
9. In contrast with purely delay-based active cancellation and variants that introduce simple gains in addition to delays (e.g., Platt & Faggin, 1992), the present invention is based on an acoustic model that includes the effects of echoes and reverberation.
10. Single-stage techniques for blind, convolutive active cancellation are usually unable to separate mixtures that are not already partially separated.
11. Two prior art techniques for blind, convolutive active cancellation are, like the present invention, based on a two-stage architecture. Of these, the technique of Najar et al. (1994) differs from the present invention in that each output channel of its first stage is a filtered version of only one input channel. Therefore, the first stage of the system described by Najar et al. (1994) cannot bring about full separation even in the absence of echoes and reverberation, unless the original sources have no overlapping spectral components.
12. The other prior art technique based on a two-stage architecture is the Griffiths-Jim beamformer (Griffiths and Jim, 1982). The Griffiths-Jim beamformer employs active cancellation in its second stage that requires a reference signal. The necessary reference noise signal is produced by the first stage, using known DOA information. If this reference noise signal contains a significant amount of the desired signal, then the second stage will erroneously enhance the noise and suppress the desired signal (Van Compernolle, 1989). In the present invention, the second stage is blind; it employs its own outputs as reference signals. Unlike the Griffiths-Jim reference signal, these become progressively purer with time as the network adapts.

What is claimed is:

1. A signal processing system for processing waves from a plurality of sources, said system comprising:

a plurality of transducer means for receiving waves from said plurality of sources, including echoes and reverberation thereof and for generating a plurality of signals in response thereto, wherein each of said plurality of transducer means receives waves from said plurality of sources including echoes and reverberations thereof, and for generating one of said plurality of signals;

first processing means for receiving said plurality of signals and for generating a plurality of first processed signals in response thereto, said first processing means comprises:

a plurality of delay means, each for receiving one of said plurality of signals and for generating a delayed signal in response thereto, and

a plurality of first combining means, each for receiving at least one of said plurality of signals and one of said delayed signals not associated with said one of said plurality of signals, and for combining said received delayed signal and said signal, by an active cancellation process, to produce one of said first processed signals; and

second processing means for receiving said plurality of first processed signals and for generating a plurality of second processed signals in response thereto, wherein each of said second processed signals represents waves from one different source, said second processing means including feedback means for supplying said plurality of second processed signals to said second processing means for combining each of said plurality of second processed signals with at least one of said plurality of first processed signals not associated with said each second processed signal to generate said plurality of second processed signals.

2. A signal processing system for processing waves from a plurality of sources, said system comprising:

a plurality of transducer means for receiving waves from said plurality of sources, including echoes and reverberation thereof and for generating a plurality of signals in response thereto, wherein each of said plurality of transducer means receives waves from said plurality of sources including echoes and reverberations thereof, and for generating one of said plurality of signals;

first processing means for receiving said plurality of signals and for generating a plurality of first processed signals in response thereto, said first processing means comprises:

a plurality of multiplying means, each for receiving different ones of said plurality of signals and for generating a scaled signal in response thereto, and

a plurality of first combining means, each for receiving at least one of said plurality of signals and one scaled signal not associated with said one of said plurality of signals, and for combining said received scaled signal and said signal to produce one of said first processed signals; and

second processing means for receiving said plurality of first processed signals and for generating a plurality of second processed signals in response thereto, wherein each of said second processed signals represents waves from one different source, said second processing me including feedback means for supplying said plurality of second processed signals to said second processing means for combining each of said plurality of second processed signals with at least one of said plurality of first processed signals not associated with said each second processed signal to generate said plurality of second processed signals.

3. The system of claim 1, further comprising:

means for generating a direction of arrival signal for said waves; and

wherein said first processing means generates said plurality of first processed signals, in response to said direction of arrival signal.

4. The system of claim 1, wherein the number of transducer means is two, the number of first processed signals is two, and the number of second processed signals is two.

5. The system of claim 1, wherein said transducer means are spaced apart omnidirectional microphones.

6. The system of claim 1 wherein said microphones are co-located directional microphones.

7. The system of claim 1, 3, 4, 5, 6, or 2 wherein said second processing means comprises:

a plurality of second combining means, each of said second combining means having a first input, at least one other input, and an output; each of said second combining means for receiving one of said first processed signals at said first input, an input signal at said other input, and for generating an output signal, at said output; said output signal being one of said plurality of second processed signals and is a difference between said first processed signal received at said first input and the sum of said input signal received at said other input;

a plurality of adaptive filter means for generating a plurality of adaptive signals, each of said adaptive filter means for receiving said output signal from one of said plurality of second combining means and for generating an adaptive signal in response thereto; and

means for supplying each of said plurality of adaptive signals to one of said other input of said plurality of second combining means other than the associated one of said second combining means.

8. The system of claim 7 further comprising means for filtering each of said second processed signals to generate a plurality of third processed signals.

9. The system of claim 8 wherein said second processed signals are characterized by having a low frequency component and a high frequency component, and wherein said filtering means boosts the low frequency component relative to the high frequency component of said second processed signals.

10. A signal processing system for processing waves from a plurality of sources, said system comprising:

a plurality of transducer means for receiving waves from said plurality of sources, including echoes and reverberations thereof and for generating a plurality of signals in response thereto, wherein each of said plurality of transducer means receives waves from said plurality of sources including echoes and reverberations thereof, and for generating one of said plurality of signals;

first processing means for receiving said plurality of signals and for generating a plurality of first processed signals in response thereto, wherein in the absence of echoes and reverberations of said waves from said plurality of sources, each of said first processed signals represents waves from only one different source; said first processing means comprising:

a plurality of delay means, each for receiving one of said plurality of signals and for generating a delayed signal in response thereto, and

a plurality of first combining means, for receiving said plurality of signals and for feedforward combining said plurality of signals in an active cancellation process to produce said plurality of processed signals, wherein each of said plurality of first combining means receives at least one of said plurality of signals and one of said delayed signals not associated with said one of said plurality of signals, and for combining said received delayed signal and said one signal to produce one of said first processed signals; and

second processing means for receiving said plurality of first processed signals and for generating a plurality of second processed signals in response thereto, wherein in the presence of echoes and reverberations of said waves from said plurality of sources, each of said second processed signals represents waves from one different source; said second processing means including feedback means for supplying said plurality of second processed signals to said second processing means for combining each of said plurality of second processed signals with at least one of said plurality of first processed signals not associated with said each second processed signal to generate said plurality of second processed signals.

11. A signal processing system for processing waves from a plurality of sources, said system comprising:

a plurality of transducer means for receiving waves from said plurality of sources, including echoes and reverberations thereof and for generating a plurality of signals in response thereto, wherein each of said plurality of transducer means receives waves from said plurality of sources including echoes and reverberations thereof, and for generating one of said plurality of signals;

first processing means for receiving said plurality of signals and for generating a plurality of first processed signals in response thereto, wherein in the absence of echoes and reverberations of said waves from said plurality of sources, each of said first processed signals represents waves from only one different source; said first processing means comprising:

a plurality of first combining means, for receiving said plurality of signals and for feedforward combining said plurality of signals in an active cancellation process to produce said plurality of processed signals,

a plurality of multiplying means, each for receiving different ones of said plurality of signals and for generating a scaled signal in response thereto; and

wherein each of said plurality of first combining means receives at least one scaled signal and one of said plurality of signals not associated with said one scaled signal, and for combining said received scaled signal and said signal to produce one of said first processed signals;

second processing means for receiving said plurality of first processed signals and for generating a plurality of second processed signals in response thereto, wherein in the presence of echoes and reverberations of said waves from said plurality of sources, each of said second processed signals represents waves from one different source; said second processing means including feedback means for supplying said plurality of second processed signals to said second processing means for combining each of said plurality of second processed signals with at least one of said plurality of first processed signals not associated with said each second processed signal to generate said plurality of second processed signals.

12. The system of claim 10 wherein said waves are acoustic waves, and said transducer means are microphones.

13. The system of claim 12 further comprising means for filtering each of said second processed signals to generate a plurality of third processed signals.

14. The system of claim 13 wherein said second processed signals are characterized by having a low frequency component and a high frequency component, and wherein said filtering means boosts the low frequency component relative to the high frequency component of said second processed signals.

15. The system of claim 10, wherein the number of transducer means is two, the number of first processed signals is two, and the number of second processed signals is two.

16. The system of claim 10, wherein said transducer means are spaced apart omnidirectional microphones.

17. The system of claim 10 wherein said microphones are co-located directional microphones.

18. The system of claim 10, 12, 13, 14, 15, 16, 17 or 11 wherein said second processing means comprises:

a plurality of second combining means, each of said second combining means having a first input, at least one other input, and an output; each of said second combining means for receiving one of said first processed signals at said first input, an input signal at said other input, and for generating an output signal, at said output; said output signal being one of said plurality of second processed signals and is a difference between said first processed signal received at said first input and the sum of said input signal received at said other input;

a plurality of adaptive filter means for generating a plurality of adaptive signals, each of said adaptive filter means for receiving said output signal from one of said plurality of second combining means and for generating an adaptive signal in response thereto; and

means for supplying each of said plurality of adaptive signals to one of said other input of said plurality of second combining means other than the associated one of said second combining means.

19. The system of claim 18 wherein each of said adaptive filter means comprises a tapped delay line.

20. A method of processing waves from a plurality of sources, comprising:

receiving said waves, including echoes and reverberations thereof, by a plurality of transducer means;

converting said waves, including echoes and reverberations thereof from said plurality of sources, by each of said plurality of transducer means into an electrical signal, thereby generating a plurality of electrical signals;

first processing said plurality of electrical signals, by an active cancellation process, to generate a plurality of first processed signals, wherein in the absence of echoes and reverberations of said waves from said plurality of sources, each of said first processed signals represents waves from one source, and a reduced amount of waves from other sources, said first processing step including:

delaying each one of said plurality of electrical signals and generating a delayed signal in response thereto, and

combining each one of said plurality of electrical signals with one of said delayed signals not associated with said one of said plurality of signals to generate one of said first processed signals; and then

secondly processing said plurality of first processed signals to generate a plurality of second processed signals, including combining each of said plurality of second processed signals with at least one of said plurality of first processed signals not associated with said each second processed signal to generate said plurality of second processed signals, wherein in the presence of echoes and reverberations of said waves from said plurality of sources, each of said second processed signals represents waves from only one different source.

21. The method of claim 20 further comprising the step of:

filtering each of said second processed signals to generate a plurality of third processed signals.

22. The method of claim 20 further comprising the step of:

sampling and converting each one of said plurality of electrical signals and for supplying same to said plurality of delay means and to said plurality of combining means, as said electrical signal.

23. The method of claim 20 wherein said second processing step further comprises:

subtracting, by a plurality of subtracting means, a different one of said first processed signals by an adaptive signal and generating an output signal, thereby generating a plurality of output signals;

adaptively filtering said output signals to generate a plurality of adaptive signals; and

supplying each one of said plurality of adaptive signals to a different one of said subtracting means.

24. A signal processing system for processing waves from a plurality of sources, said system comprising:

at least a first and second transducer for receiving waves from said plurality of sources, including echoes and reverberation thereof and for generating at least a first and a second signal in response thereto, wherein each of said transducers receives waves from said plurality of sources including echoes and reverberations thereof, and for generating one of said first and second signals;

first processing means for receiving said first and second signals and for generating a first and a second processed signals in response thereto, said first processing means comprises:

first delay means for receiving said first signal and for generating a first delayed signal in response thereto,

second delay means for receiving said second signal and for generating a second delayed signal in response thereto,

first combining means for receiving said first signal and said second delayed signal, and for combining said received first signal and said second delay signal, by an active cancellation process, to produce said first processed signal, and

second combining means for receiving said second signal and said first delayed signal, and for combining said received second signal and said first delayed signal, by an active cancellation process, to produce said second processed signal; and

second processing means for receiving said first and second processed signals and for generating a third and a fourth processed signals in response thereto, said second processing means comprises:

third combining means for receiving the first processed signal to produce the third processed signal in response thereto;

fourth combining means for receiving the second processed signal to produce the fourth processed signal in response thereto;

first adaptive filter means for receiving said third processed signal, for generating a first adaptive signal in response thereto, and for supplying said first adaptive signal to said fourth combining means;

second adaptive filter means for receiving said fourth processed signal, for generating a second adaptive signal in response thereto, and for supplying said second adaptive signal to said third combining means;

wherein the third combining means combines the first processed signal and the second adaptive signal to produce the third processed signal so that the third processed signal is a difference between the first processed signal and the second adaptive signal; and

wherein the fourth combining means combines the second processed signal and the first adaptive signal to produce the fourth processed signal so that the fourth processed signal is a difference between the second processed signal and the first adaptive signal.

* * * * *

United States Patent [19]

Lea

[11] **Patent Number: 4,601,025**

[45] **Date of Patent: Jul. 15, 1986**

[54] **ANGLE TRACKING SYSTEM**

[75] Inventor: **John D. Lea**, Huntington, N.Y.

[73] Assignee: **Sperry Corporation**, New York, N.Y.

[21] Appl. No.: **546,758**

[22] Filed: **Oct. 28, 1983**

[51] **Int. Cl.**[4] **G01S 3/82**

[52] **U.S. Cl.** **367/125**; 367/127

[58] **Field of Search** 367/125, 127; 343/378

[56] **References Cited**

U.S. PATENT DOCUMENTS

3,249,911	5/1966	Gustafsson	367/125
3,545,000	12/1970	Heflinger	367/127 X

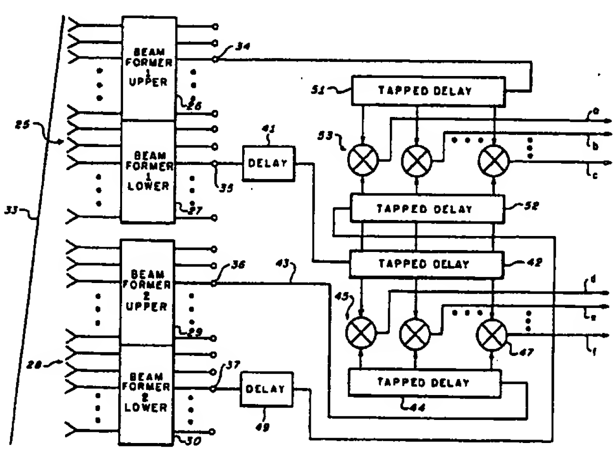
Primary Examiner—Richard A. Farley

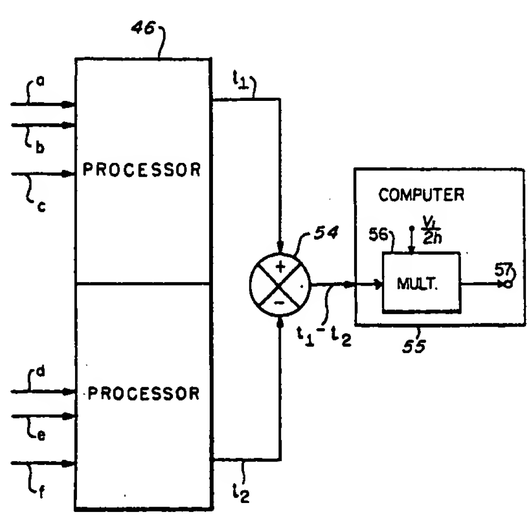
Attorney, Agent, or Firm—Howard P. Terry; Seymour Levine

[57] **ABSTRACT**

An apparatus for determining an angle of incidence of a received signal utilizes two widely spaced interferometers with equal length crossing baselines. The time differential of a signal arrival between the two elements of each interferometer is established and the difference between the two time differentials is determined. Multiplying this difference by a predetermined constant provides the desired angle.

9 Claims, 9 Drawing Figures





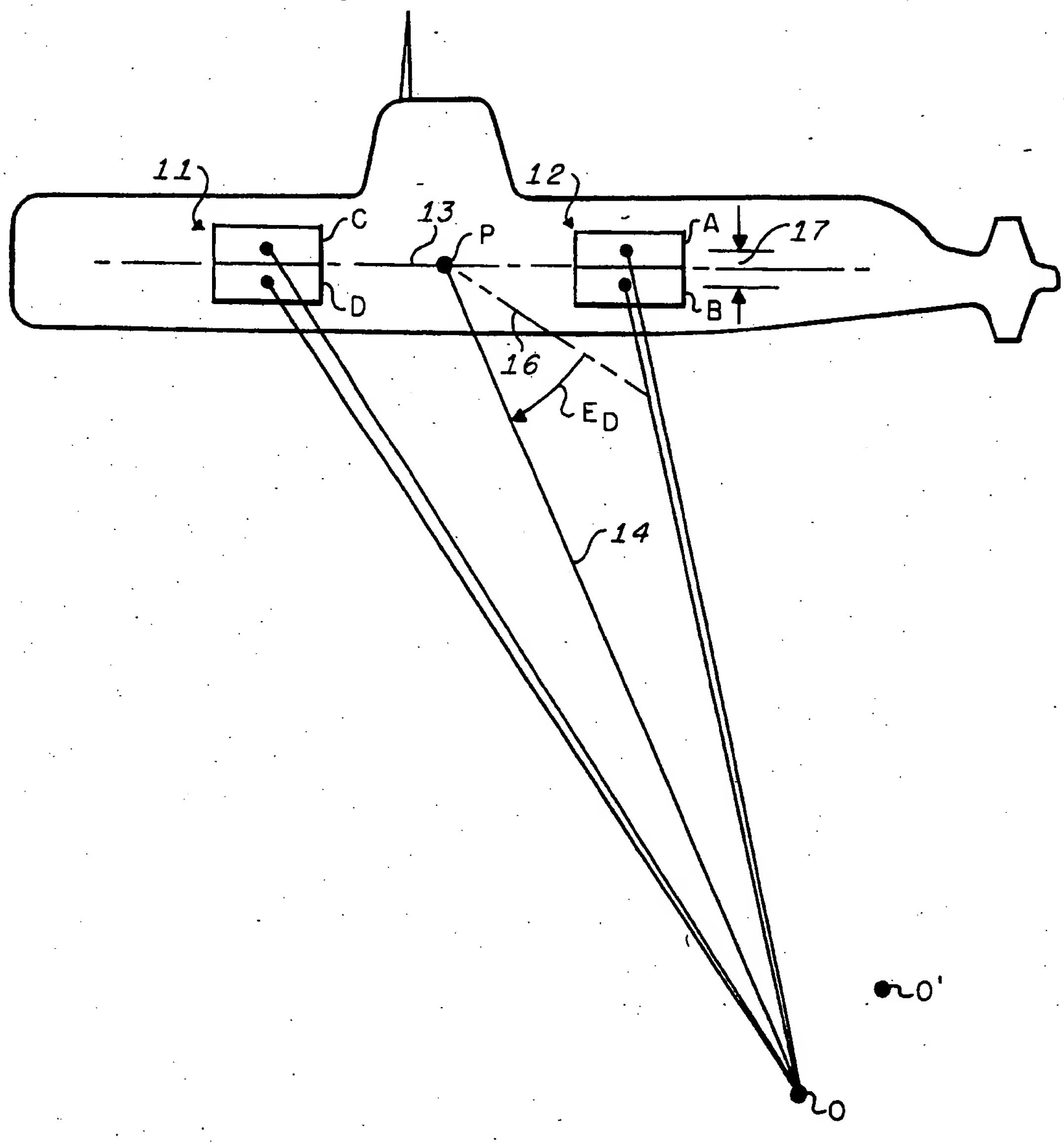


FIG.1.

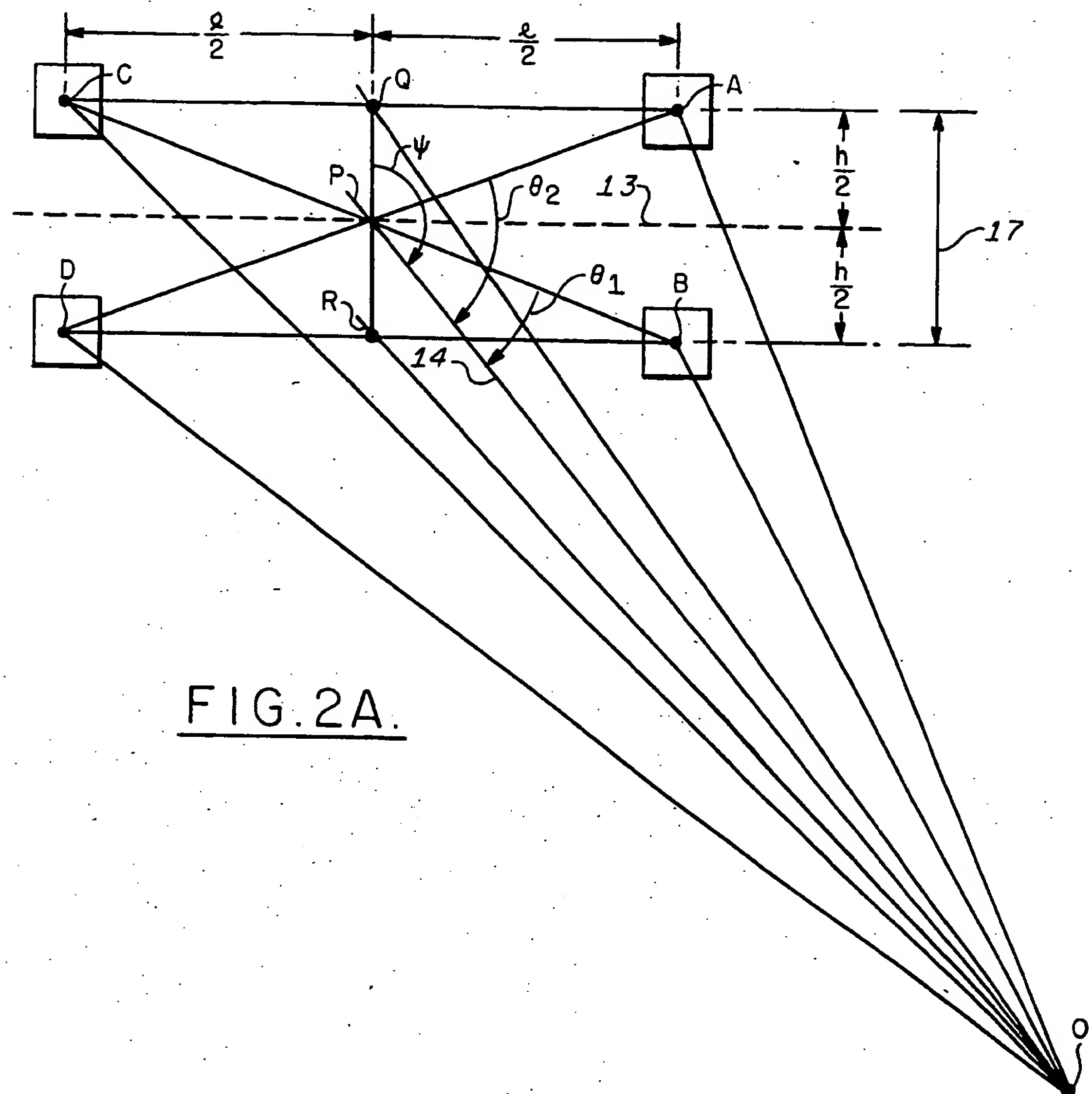

C
Q
A
P
ψ
θ_2
13
17
$\frac{l}{2}$
$\frac{h}{2}$
D
R
B
θ_1
14
O

FIG. 2A.

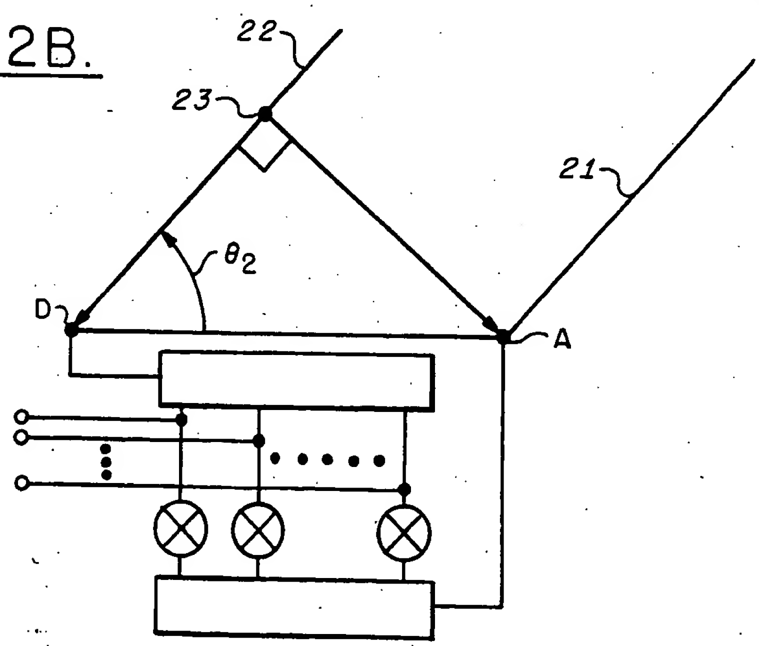

22
23
21
θ_2
D
A

FIG. 2B.

(1) $$\cos\psi = \frac{\left(\frac{h}{2}\right)^2 + (\overline{OP})^2 - (\overline{OQ})^2}{h\,(\overline{OP})}$$

(2) $$-\cos\psi = \frac{\left(\frac{h}{2}\right)^2 + (\overline{OP})^2 - (\overline{OR})^2}{h\,(\overline{OP})}$$

(3) $$\cos\psi = \frac{(\overline{OR})^2 - (\overline{OQ})^2}{2h\,(\overline{OP})}$$

(4) $$(\overline{OQ})^2 = \frac{1}{2}\left[(\overline{OA})^2 + (\overline{OC})^2 - \frac{\ell^2}{2}\right]$$

(5) $$(\overline{OR})^2 = \frac{1}{2}\left[(\overline{OD})^2 + (\overline{OB})^2 - \frac{\ell^2}{2}\right]$$

(6) $$\cos\psi = \frac{\left[(OD)^2 - (OA)^2\right] + \left[(OB)^2 - (OC)^2\right]}{4\,(\overline{OP})\,h}$$

(7a) $$(\overline{OD})^2 - (\overline{OA})^2 = 2(\overline{AD})(\overline{OP})\cos\theta_2$$

(7b) $$(\overline{OB})^2 - (\overline{OC})^2 = -2(\overline{BC})(\overline{OP})\cos\theta_1$$

(7c) $$\overline{AD} = \overline{BC} = L$$

(8) $$\cos\psi = \frac{\overline{AD}}{2h}\left[\cos\theta_2 - \cos\theta_1\right]$$

(9a) $$\cos\theta_1 = \frac{V_1}{L}t_1$$ (9b) $$\cos\theta_2 = \frac{V_1}{L}t_2$$

(10) $$\cos\psi = \frac{V_1}{2h}(t_2 - t_1)$$

(11) $$E_D = \psi - 90°$$

(12) $$\sin E_D = \frac{V_1}{2h}(t_1 - t_2)$$

(13) $$t_D = \frac{L}{V_1}\cos\theta_i$$

(14) $$S = 2\frac{V_2}{V_1}L\sin\frac{\theta_B}{4}\sin\left(\theta_i + \frac{\theta_B}{4}\right)$$

(15) $$X = \frac{S}{2} + \frac{V_2}{V_1}L\sin\frac{\theta'}{2}\sin\left(\theta_i + \frac{\theta'}{2}\right)$$

(16) $$t_1 - t_2 = \frac{2(X_1 - X_2)}{\nu_s}$$

(17) $$\Phi_{11}(\tau) = \frac{1}{T}\int_{-\frac{T}{2}}^{T/2} f(t)f(t+\tau)\,dt$$

FIG. 3.

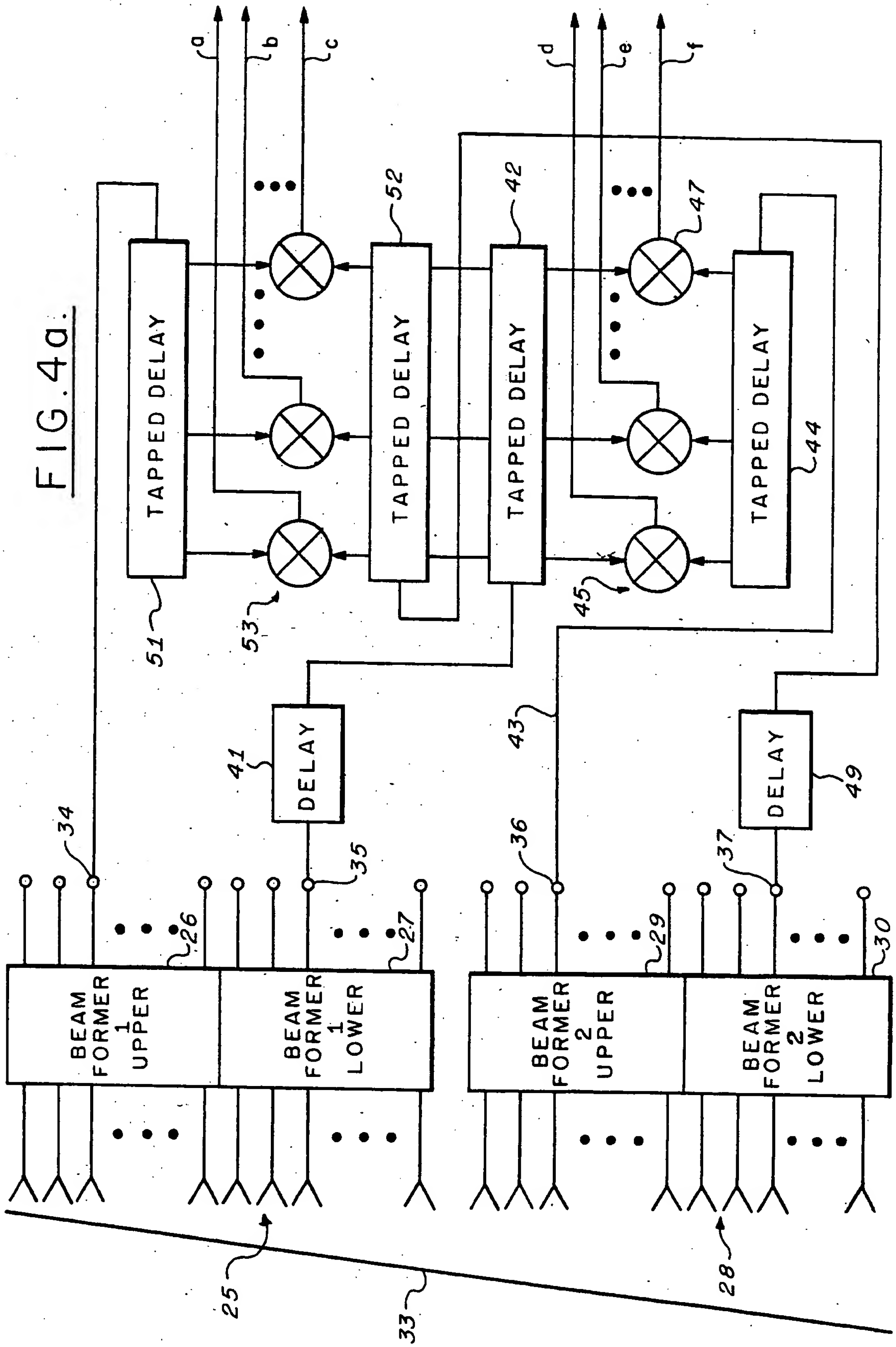
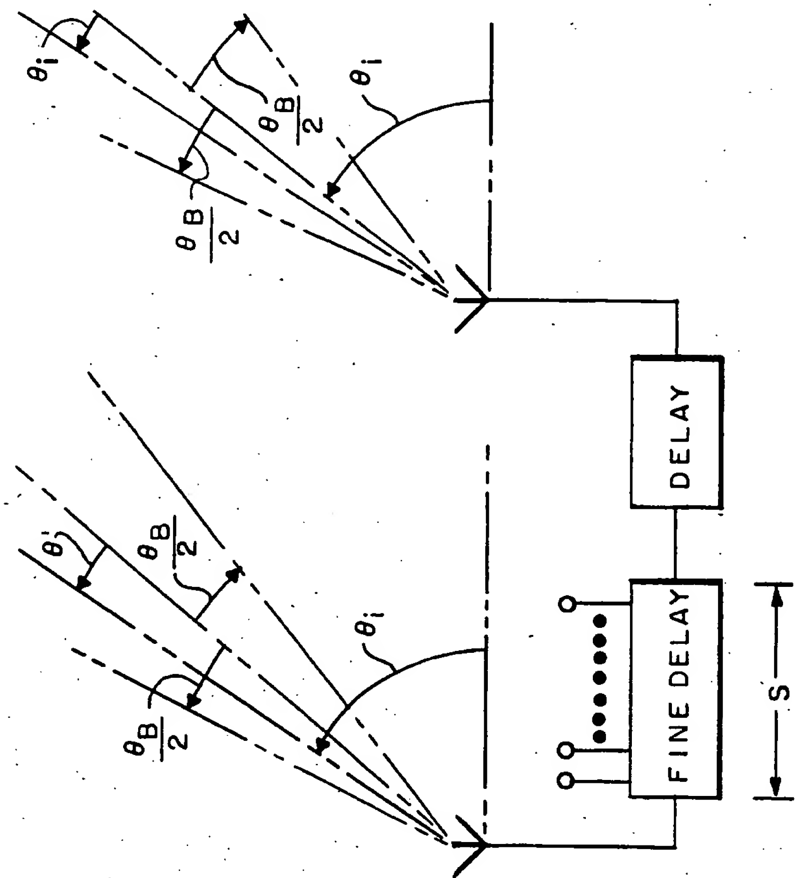

a
b
c
d
e
f
TAPPED DELAY
TAPPED DELAY
TAPPED DELAY
TAPPED DELAY
DELAY
DELAY
BEAM FORMER 1 UPPER
BEAM FORMER 1 LOWER
BEAM FORMER 2 UPPER
BEAM FORMER 2 LOWER
25
26
27
28
29
30
33
34
35
36
37
41
42
43
44
45
47
49
51
52
53

FIG. 4a.

FIG. 5.


DELAY
FINE DELAY
S
θ_i
θ_i'
$\frac{\theta_B}{2}$

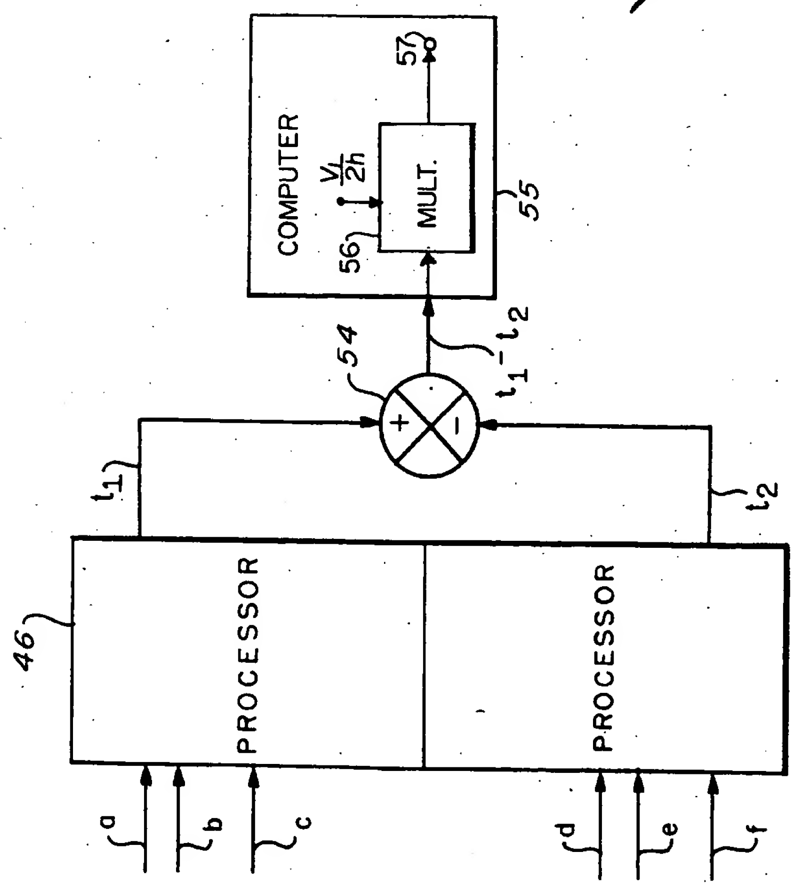

PROCESSOR
PROCESSOR
46
a
b
c
d
e
f
t_1
t_2
54
$t_1 - t_2$
COMPUTER
MULT.
$\frac{V_L}{2h}$
55
56
57

FIG. 4b.

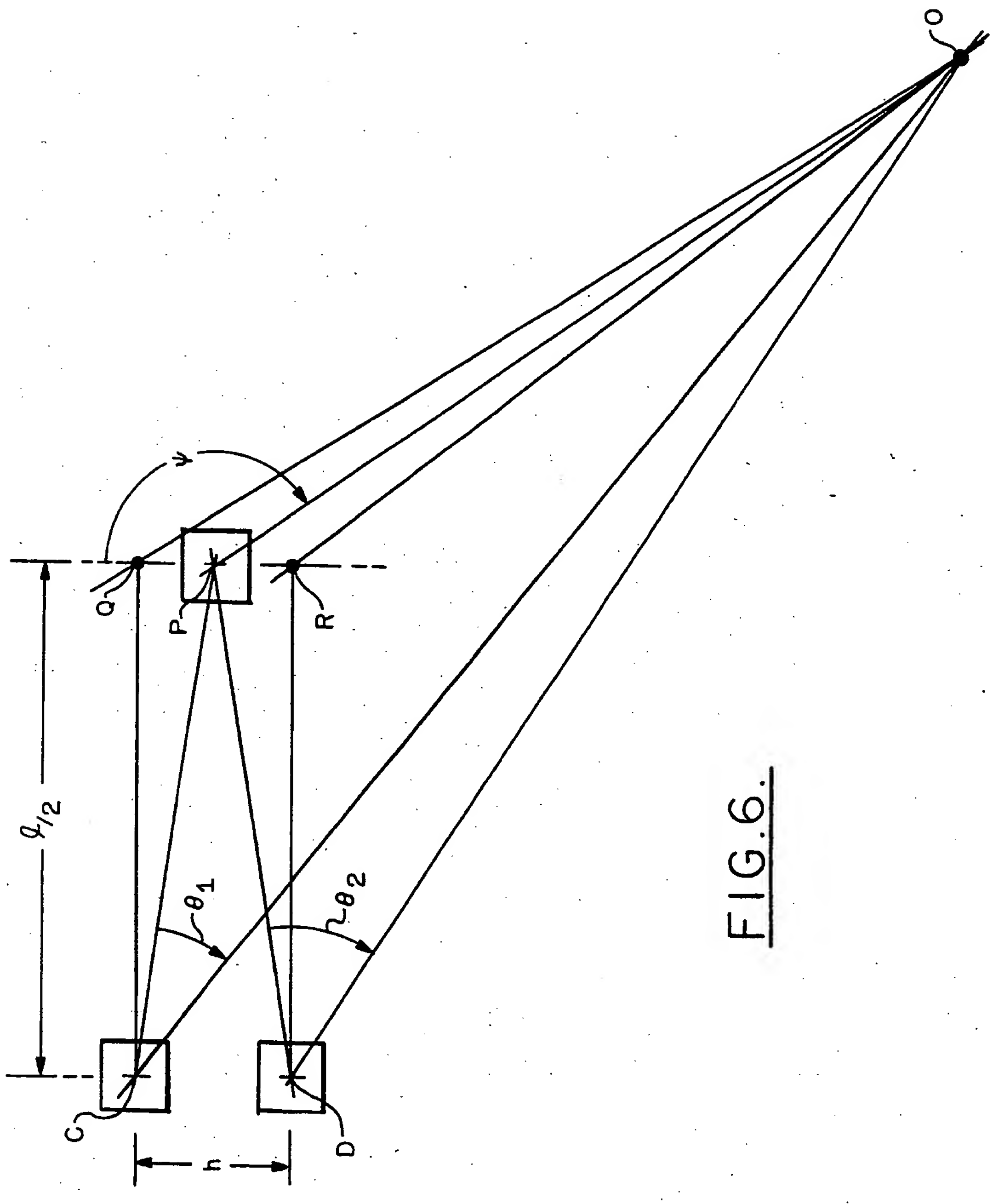


FIG. 6.

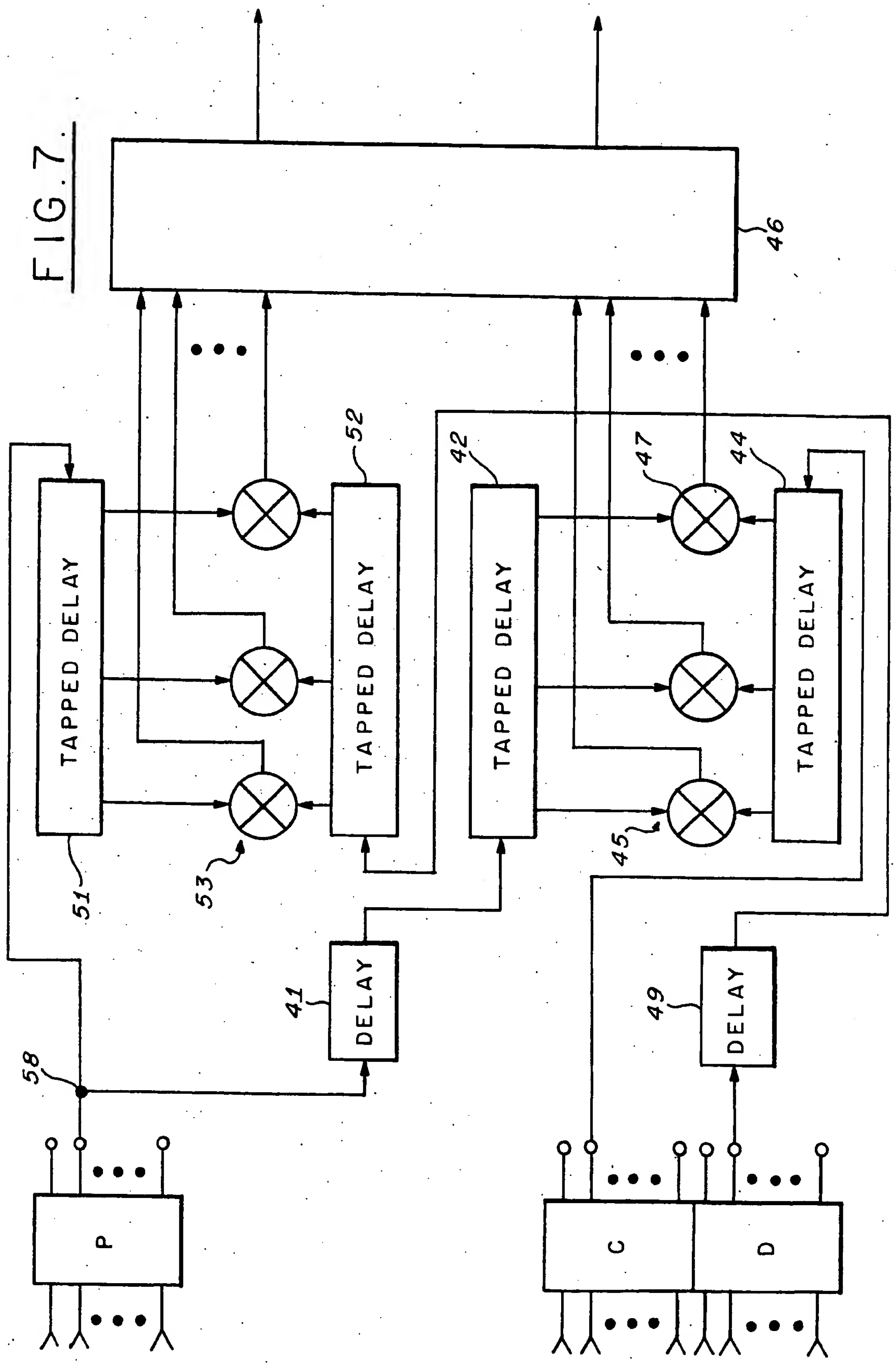

46
TAPPED DELAY
TAPPED DELAY
TAPPED DELAY
TAPPED DELAY
DELAY
DELAY
51
52
53
42
47
44
45
41
49
58
P
C
D

FIG. 7.

ANGLE TRACKING SYSTEM

BACKGROUND OF THE INVENTION

1. Field of the Invention

The invention pertains to angle of signal arrival measurements and more specifically to the measurement of the vertical angle of the signal arriving from a source not necessarily in the horizontal plane of the receiver.

2. Description of the Prior Art

Interferometric methods for measuring arrival angles of received signals have long been in use in the sonar and radar art. In these systems, the time difference of arrival at receivers positioned a predetermined distance apart is determined by correlation techniques, if the signals are continuous, or by timing techniques, if the signals are pulses. This difference in time of arrival and the separation distance is utilized to determine the angle of arrival of the signal with an accuracy that is a function of the receiver separation, improving as the separation increases. The measured angle is in the plane of signal propagation defined by the two receivers and the source. In many applications, however, the horizontal plane source angle i.e. azimuth or bearing is required rather than the angle in the signal propagation plane provided by the interferometer. A conversion from the signal propagation angle to the horizontal plane angle may readily be realized with knowledge of the vertical angle of arrival of the signal. Vertical angle of arrival is also useful for other applications, such as establishing the relative altitude or depth of the signal source. Prior art methods of measuring vertical angle are of limited accuracy and resolution capability.

Vertical angle measurements have been made in the prior art with two beams having peaks offset at equal and opposite angles from a reference angle to establish equal amplitude responses for signals incident from the reference direction. In a sonar system, an acoustic signal arriving from the reference angle direction induces electrical signals in the beam transducers of equal magnitude establishing a zero signal difference therebetween. Acoustic signals arriving from angular directions other than the reference angle induce electrical signals that differ from zero, having a magnitude as a function of the angle off the reference angle and a polarity which is determined by whether the arrival angle is less than or greater than the reference angle. The accuracy of these systems is generally poor, being a function of the relative beam shapes of the transducers and the interpolation.

Greater accuracy than that obtainable with amplitude interpolation systems may be achieved with the utilization of a vertically split array to effectively establish two transducer arrays with a physical separation therebetween. These systems determine the time difference of arrival of a signal incident to the dual array. This time difference is a function of the angle from the perpendicular to the array surface and does not depend on the array beam shape, being only a function of the angle of arrival and the dual array separation. Array separation, however, is generally small requiring that the difference between two nearly equal times of arrival be determined, thus limiting the accuracy of the system. Additionally, the small separation between the dual arrays causes the noise at the output terminals of each array to be correlated, adversely affecting the signal-to-noise ratio of the system and concomitantly the angle determination accuracy. Further, the dual arrays overlap in the azimuthal plane and thereby cannot resolve targets within the azimuthal beam width of the arrays. This limitation of bearing resolution causes the dual array to respond to the centroid of multiple targets within the aximuthal beam width.

SUMMARY OF THE INVENTION

An angle measuring system constructed in accordance with the principles of the present invention includes two widely separated split array pairs, each pair comprising upper and lower arrays with parallelly scanned receiving beams. The upper array of one pair and the lower array of the other are coupled to form two wide based interferometer pairs. In each pair the signal arrival time differential between the arrays is determined and utilized to establish the desired angle to the target.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a pictorial representation of dual arrays mounted on a ship with ray lines to the target indicated thereon.

FIGS. 2A and 2B depict the ray line geometry between a target of interest and the dual split arrays.

FIG. 3 is a tabulation of formulas useful for explaining the invention.

Figure numeral 4*a* and 4*b* comprise a block diagram of an embodiment of the invention.

FIG. 5 is a geometrical representation of ray paths within an array beamwidth that is useful in the determination of the length of the delay line employed in the invention.

FIG. 6 depicts the ray line geometry of a three receiver implementation of the invention.

FIG. 7 is a block diagram of another embodiment of the invention.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

In FIG. 1, sonar arrays **11**, **12** are shown mounted on a ship along a wide baseline **13** which is coincident with a horizontal reference axis. Each array is split to form two subarrays giving rise to subarrays A, B, C, D. Range **14** to the target is measured from the midpoint P between the arrays while the depression angle E_D is measured from the horizontal **16**. This depression angle may be determined from the separation **17** between the phase centers of the split array and the deferential time of arrival therebetween of a signal emitted from the target O. Since the separations of the phase centers of the split arrays A, B and C, D is short, the received noise in adjacent split arrays is correlated, as for example the noise received by the split arrays A, B. This correlation of received noise adversely affects the depression angle measurement accuracy. Additionally, when the signals received by the four arrays A, D, C, B from the target O are processed as in the prior art, signals from targets at O and O' that are within an array beamwidth cannot be resolved and angular measurements corresponding to the target signal centroid result. These deficiencies of the prior art are remedied by determining the differential time of arrival of received signals at subarrays A, D and subarrays C, B, as will be described subsequently. The long baseline established between the correlating subarrays significantly reduces the noise signal correlation at the subarrays and signifi-

cantly improves the target resolution capability of the system.

Refer now to FIG. 2A wherein previously referenced elements are given the previously assigned reference numerals and wherein line RQ is coincident with a vertical reference axis. Consider the triangle OPQ, where Q is the midpoint between the phase centers of the subarrays A, C and is h/2 vertically distant from the midpoint of the overall array P, h being the vertical subarray phase center separation 17. The horizontal subarray phase center separation being l. The application of the law of cosines to the triangle OPQ yields the equation 1 of FIG. 3 for the cosine of the angle ψ between the vertical PQ and the range ray OP. Similarly, the application of the law of cosines to the triangle OPR, where R is the midpoint between the subarrays B, D yields the expression for cos ψ given by equation 2. Equation 3 is an expression for cos ψ in terms of the distances OR, OQ, and OP that is obtained by subtracting equation 2 from equation 1. Expressions for distances OQ and OR shown in equations 4 and 5 may be obtained by applying the law of cosines to triangles OAC, OAQ and OBD, OBR, respectively. Substituting equations 4 and 5 into equation 3 yields the expression for cos ψ given in equation 6. Expressions for the bracketed terms in the numerator of equation 6 may be obtained by applying the law of cosines to the triangles in the planes OAD and OBC. These expressions are shown in equations 7*a* and 7*b*. Equation 7*c* states that the diagonal distance AD is equal to the diagonal distance BC. This is a consequence of the positioning of the phase centers of the subarrays at the corners of a rectangle. The substitutions of the equations 7 into equation 6 yields an expression for cos ψ in terms of the angles of incidence to the diagonally positioned interferometers A, D and C, B are shown in equation 8.

Refer now to FIG. 2B wherein ray 21 and ray 22 are respectively shown incident to subarray A and subarray D. Rays 21, 22 correspond to the rays OA and OD of FIG. 2A, respectively, when the position O is at a distance from the array that is very much greater than the length of the baseline AD. Under these conditions, the rays 21, 22 are substantially parallel and the path length difference to the subarrays A, D of signals emitted from a target at position O may be determined by dropping a perpendicular from the subarray A to intercept ray 22 at a point 23. The distance between point 23 and subarray D is therefore the differential path length to subarrays A,D of a signal emitted from a target of position O. Since the rays 21, 22 are substantially parallel, it follows that the ray OP of FIG. 2A is substantially parallel to these rays and the angle between AD and D (23) is equal to the angle θ_2 shown in FIG. 2A. Similarly, the differential ray paths of a signal emitted by a target at position O to the phase centers of subarrays B, C may be determined from the baseline BC and the angle θ_1. From the above it is apparent that cos θ_1 and cos θ_2 are given by equations 9*a*, 9*b*, wherein V_1 is the propagation velocity of the wavefront 33 and t_1 and t_2 are the differential times of arrival. Consequently, the difference between the differential time delays of the interferometers AD, BC determines the depression angle ψ from the vertical as given in equation 10. Since the depression angle from the horizontal E_D is equal to the depression angle from the vertical less 90 degrees as shown in equation 11, it follows that depression angle from the vertical E_D is given by equation 12.

Referring now to FIG. 4, a first array of receiving elements 25 may be coupled to beam formers 26, 27 to form upper and lower receiving beams and a second array of elements 28 may be coupled to beam formers 29, 30 to form a second set of upper and lower receiving beams. A phase front 33 incident to arrays 25, 28 induces a signal at output port 34 of beam former 26, output port 35 of beam former 27, output port 36 of beam former 29, and output port 37 of beam former 30. Each of the output ports couple to signals that arrive at angles within a beam centered at an angle θ_i from the baseline between the phase centers of an upper and lower beam combination, as for example the phase center of the array of elements coupled to upper beam former 26 and the phase center of the array of elements coupled to lower beam former 30, as represented in FIG. 5. Signals at the output port 35 of lower beam former 27 are coupled through a delay line 41 to the input terminal of a tapped delay line 42, while signals at output terminal 36 of upper beam former 29 are coupled via line 43 to the input terminal of a tapped delay line 44. Each tap on the tapped delay line 44 is an incremental fine time delay from the coarse time delay of delay line 41. The input terminals to delay lines 42, 44 are oppositely positioned to provide differential time delays as will be explained subsequently. Corresponding output terminals of delay lines 42, 44 are coupled to correlators 45 wherein signals at corresponding taps are correlated and wherefrom a multiplicity of correlation signals is coupled to a processor 46.

Delay line 41 provides a delay t_D that is given by equation 13 in FIG. 3, L being the length of the baseline. When a signal arrives at an angle θ_i corresponding to the beam peak, the delayed output signal from terminal 35 and the undelayed output signal from terminal 36 arrive at the input terminals of delay lines 42, 44 at substantially the same time and are thereby in phase at the central taps of the delay lines 42, 44, thus establishing a correlation output signal from the correlator coupled to the central tap that exceeds all output signals from the correlators coupled to the other taps of the delay lines. If the signal is incident at an angle other than that corresponding to the beam peak, the signals arrive at the input terminals at delay lines 42, 44 with a differential time lag therebetween. This differential time lag causes the peak correlation signal to appear at a tap on either side of the central tap. The tap at which this peak appears is a function of the incident angle to the arrays 25, 28. If the beam width at the output terminals 35, 36 is θ_B as shown in FIG. 5 and the length S of the tap delay lines is chosen such that a signal arriving at an angle $\theta_i - \theta_{B/2}$ establishes a maximum correlation signal at the correlator 47 coupled to the last tap of the delay line 42 and the first tap of delay line 44, S is then given by equation 14 of FIG. 3 wherein V_2 is the propogation velocity along the tapped delay lines. For an angle θ' within the beam the maximum correlations will occur at a distance from the input end of tap delay line 43 that is given by equation 15. The output signals from the correlators 45 are coupled to processor 46, which may contain a network of comparators and logic circuits, to determine the taps that give rise to the maximum correlation signal, thereby determining the time differential of the received signals at the phase centers of the interferometer formed by the upper beam former 29 and the lower beam former 27. Similarly, the time differentials of received signals at the phase centers of the upper beam former 26 and the lower beam former 30 are de-

termined by the delay line **49**, tap delay lines **51**, **52** and correlators **53**. Time differentials so determined are coupled to substraction network **54** wherefrom the difference in the time differentials, given by equation **16**, is coupled to a computer **55** for the determination of a depression angle in accordance with equation **12**. Computer **55** contains a multiplier **56** to which signals representative of $V_1/2h$ and (t_1-t_2) are coupled for multiplication to provide a signal at an output terminal **57** that is representative of the depression angle E_D.

The signals coupled to each of the correlators **45**, **53** emanate from the same source but arrive at the correlators with time delay differentials that are functions of the angles of incidence and the time delays of the system. Thus, each correlator performs an autocorrelation in accordance with the well known formula **17** of FIG. **3**. Autocorrelation function equation **17** provides a peak value when the differential time delay τ is zero. Thus, the correlator that provides the maximum correlation signal is that for which the signals coupled thereto arrive in phase. It should be recognized by those skilled in the art that the correlators at which the correlation signals are maximum are equally and oppositely displaced from the central correlator in correlators **45**, **53**. For each target within a subarray beam, an additional pair of peak correlations are induced in the correlators **45**, **53**. Since the peaks of each pair are equally and oppositely spaced from the central correlator, each pair may be identified and appropriately processed as described above. In this manner, multiple targets within a subarray beam, but at varying depression angles may be resolved and tracked.

It should be recognized that the above-described angle determination may also be accomplished by positioning a receiver at point P, the crossover of the two base lines $\overline{AD}$, $\overline{CB}$, to form a triangle with receivers at points C and D, as shown in FIG. **6**. The time differences of arrival t_1' and t_2' measured between the receiver at P and the receivers at C and D are one-half the time differences of arrival t_1 and t_2, respectively. With the substitution $t_1=2t_1'$ and $t_2=2t_2'$ the equations of FIG. **3** may be utilized to determine the vertical angle ψ. The measurement of the time of arrival may be accomplished in the manner previously described.

Referring to FIG. **7**, wherein elements previously discussed bear the prior assigned reference numerals, the output signals of the receiver at C may be coupled to the tapped delay line **44** while the output signals of the receiver at D may be coupled via coarse delay line **49** to tapped delay line **52**. Output signals of the receiver at P are coupled from node **58** directly to tapped delay line **51** and via coarse delay line **41** to tapped delay line **42**.

While the invention has been described in its preferred embodiments, it is to be understood that the words which have been used are words of description rather than limitation and that changes may be made within the purview of the appended claims without departing from the true scope and spirit of the invention in its broader aspects.

I claim:

1. A passive apparatus for measuring an angle to a signal emitter comprising:
   - first and second means positioned with a predetermined separation distance along a first axis therebetween for receiving signals emitted from said emitter signal;

   - third and fourth means positioned with said predetermined separation distance along said first axis for receiving said emitted signals, located a preselected distance from said first and second means along a second axis and relatively positioned along said first axis such that said first and third means and said second and fourth means are correspondingly positioned along said first axis;
   - time difference means coupled to said four receiving means for providing a signal representative of a first differential time t_1 between a signal arrival at said first receiving means and said signal arrival at said fourth receiving means and a signal representative of a second differential time t_2 between said signal arrival at said second receiving means and said third receiving means: and
   - means for providing a signal representative of a difference between times t_1 and t_2, said difference between t_1 and t_2 being representative of said angle to said signal emitter.

2. A passive angle measuring apparatus in accordance with claim **1** wherein said time difference means includes:
   - delay means coupled to receive signals from said second and fourth receiving means for providing time delays substantially equal to time differentials of arrival at predetermined angles of incidence between said first and fourth and said third and second receiving means;
   - fine delay means coupled to said delay means to receive delayed signals from said second and fourth receiving means and to receive signals directly from said first and third receiving means for providing a multiplicity of fine time delays to said delayed signals and to said directly received signals; and
   - means coupled to said fine delay means to receive said delayed signals and said directly received signals in pairs, after corresponding fine time delays, for correlating said fine time delayed pairs of delayed signals and directly received signals.

3. A passive receiving apparatus for angular measurements in accordance with claim **2**, wherein said fine delay means includes:
   - first tapped delay line means for providing said fine time delays having input means at ends thereof coupled to receive signals directly from said first and third receiving means; and
   - second tapped delay means for providing said fine time delays having input means at ends thereof opposite said receiving ends of said first tapped delay means coupled to receive signals from said second and fourth receiving means delayed through said delay means;
   - said first and second tapped delay means having taps correspondingly positioned to form tap pairs, each pair of taps coupled to said correlating means.

4. A method of measuring an angle to a signal emitter, which comprises:
   - receiving a signal from said signal emitter at first, second, third, and fourth receiving means, positioned with phase centers at corners of a predetermined rectangle, said first and fourth receiving means and said second and third receiving means forming diagonally positioned pairs;
   - delaying signals received by said second and fourth receiving means for a time duration determined by said rectangular diagonal length and a signal angle

of incidence corresponding to a peak of a receiving beam;

coupling said delayed signals from said second receiving means and signals from said third receiving means to first and second fine delay means, respectively, said first and second fine delay means having a multiplicity of corresponding output terminal pairs, each pair representative of an angle within said receiving beam;

coupling said delayed signals from said fourth receiving means and signals from said first receiving means to third and fourth fine delay means, respectively, said third and fourth fine delay means having a multiplicity of corresponding output terminal pairs, each pair representative of an angle within said receiving beam;

correlating signals at said terminal pairs of said first and second fine delay means and said third and fourth fine delay means;

determining output terminal pairs at which peak correlation signals occur; and

establishing angle of signal incidence from said terminal pair peak correlation signal determination.

5. The method of claim 4, wherein the step of establishing said angle of signal incidence includes:

determining the differential time of arrival of said incident signal between said first and fourth receiving means and between said second and third receiving means;

subtracting the differential time of arrival between said first and fourth receiving means from said differential time of arrival between said second and third receiving means; and

multiplying said difference between said differential time delays by a predetermined constant factor.

6. A passive apparatus for measuring an angle to a signal emitter comprising:

first and second means positioned on a first axis with a predetermined separation distance therebetween for receiving signals emitted from said signal emitter;

third means located a preselected distance from said first and second means along a second axis for receiving signals emitted from said signal emitter;

delay means coupled to receive signals from said first and third receiving means for providing time delays substantially equal to time differentials of arrival at predetermined angles of incidence between said first and third and said second and third receiving means;

fine delay means coupled to said delay means to receive delayed signals from said first and third receiving means and to receive signals directly from said second and third receiving means for providing a multiplicity of time delays to said delayed signals and to said directly received signals, said time delays and said fine time delays providing a first differential time t_1 between a signal arrival at said first receiving means and a signal arrival at said third receiving means, and a second differential time t_2 between a signal arrival at said second receiving means and a signal arrival at said third receiving means;

means coupled to said fine delay means to receive said delayed signals and coupled to receive said directly received signals, after corresponding fine

time delays, for correlating said fine time delayed signals and directly received signals; and

means for providing a signal representative of difference between times t_1 and t_2 that is representative of said angle to said signal emitter.

7. A passive receiving apparatus for angular measurements in accordance with claim 6, wherein said fine delay means includes:

first tapped delay line means for providing said fine time delays having input means at ends thereof coupled to receive signals directly from said second and third receiving means; and

second tapped delay means for providing said fine time delays having input means at ends thereof opposite said receiving ends of said first tapped delay means coupled to receive signals from said first and third receiving means delayed through said delay means;

said first and second tapped delay means having taps correspondingly positioned to form tap pairs, each pair of taps coupled to said correlating means.

8. A method of measuring an angle to a signal emitter, which comprises:

receiving a signal from said signal emiter at first, second and third receiving means, positioned with phase centers at corners of a predetermined isoceles triangle with a base through said phase centers of said first and second receiving means;

delaying signals received by said first and third receiving means for a time duration determned by said isoceles triangle side length and a signal angle of incidence corresponding to a peak of a beam incident to said receiving means:

coupling said delayed signals from said first and third receiving means to first and second fine delay means, respectively, said first and second fine delay means having a multiplicity of output terminals;

coupling said signals from said second and third receiving means directly to third and fourth fine delay means, respectively, said third and fourth fine delay means having a multiplicity of output terminals respectively corresponding to output terminals of said first and second fine delay means to form terminal pairs, each pair representative of an angle within said beam incident to said receiving means;

correlating signals at said terminal pairs of said first and third fine delay means and said second and fourth fine delay means;

determining output terminal pairs at which peak correlation signals occur; and

establishing angle of signal incidence from said terminal pair peak correlation signal determination.

9. The method of claim 8, wherein the step of establishing said angle of signal incidence includes:

determining the differential time of arrival of said incident signal between said first and third receiving means and between said second and third receiving means;

subtracting the differential time of arrival between said first and third receiving means from said differential time of arrival between said second and third receiving means; and

multiplying said difference between said differential time delays by a predetermined constant factor.

* * * * *

US005581620A

United States Patent [19]

Brandstein et al.

[11] **Patent Number: 5,581,620**

[45] **Date of Patent: Dec. 3, 1996**

[54] **METHODS AND APPARATUS FOR ADAPTIVE BEAMFORMING**

[75] Inventors: **Michael S. Brandstein**, Providence; **Harvey F. Silverman**, East Greenwich, both of R.I.

[73] Assignee: **Brown University Research Foundation**, Providence, R.I.

[21] Appl. No.: **231,646**

[22] Filed: **Apr. 21, 1994**

[51] **Int. Cl.**[6] **H04R 3/00**

[52] **U.S. Cl.** **381/92**; 367/125; 367/126

[58] **Field of Search** 381/122, 92, 66, 381/26, 155; 367/125, 124, 126, 121, 123

[56] **References Cited**

U.S. PATENT DOCUMENTS

3,579,104	5/1971	Pignard et al.	367/125
3,787,849	1/1974	Sletten et al. .	
4,017,859	4/1977	Medwin .	
4,112,430	9/1978	Ladstatter .	
4,131,760	12/1978	Christensen et al.	381/66
4,333,170	6/1982	Mathews et al.	367/125
4,639,733	1/1987	King et al. .	
4,641,143	2/1987	Apostolos .	
4,651,155	3/1987	Baurle et al. .	
4,658,426	4/1987	Chabries et al. .	
4,723,292	2/1988	Taguchi .	
4,741,038	4/1988	Elko et al.	381/92
4,752,961	6/1988	Kahn	381/92
4,754,282	6/1988	Edelbluate et al. .	
4,769,847	9/1988	Taguchi .	
4,811,404	3/1989	Vilmur et al. .	
5,218,359	6/1993	Minamisono .	

FOREIGN PATENT DOCUMENTS

352537A1	2/1985	Germany .

OTHER PUBLICATIONS

Zibulski et al., "Oversampling in the Gabor Scheme," IEEE, Transaction Signal Processing pp. 281–284 (1992).

Flanagan et al., "Computer–Steered Microphone Arrays for Sound Transduction in Large Room," Journal Of Acoustical Socieity Of America, vol. 78, No. 5, pp. 1508–1518 (1985).

Greenberg et al., "Evaluation of an Adaptive Beamforming Method for Hearing Aids," Journal Of Acoustical Society Of America, vol. 91, No. 3, pp. 1662–1676 (1992).

Flanagan, "Bandwidth Design for Speech–Seeking Microphone Arrays," Proceedings Of 1985 ICASSP, pp. 732–735, (1985).

Kellermann, "A Self–Steering Digital Microphone Array," Proceedings Of 1991 ICASSP, pp. 3581–3584 (1991).

Yau et al., "Image Restoration by Complexity Regularization Via Dynamic Programming," IEEE, International Conference On Acoustic Speech, and Signal Processing, vol. 3, pp. 305–308 (1992).

(List continued on next page.)

Primary Examiner—Curtis Kuntz

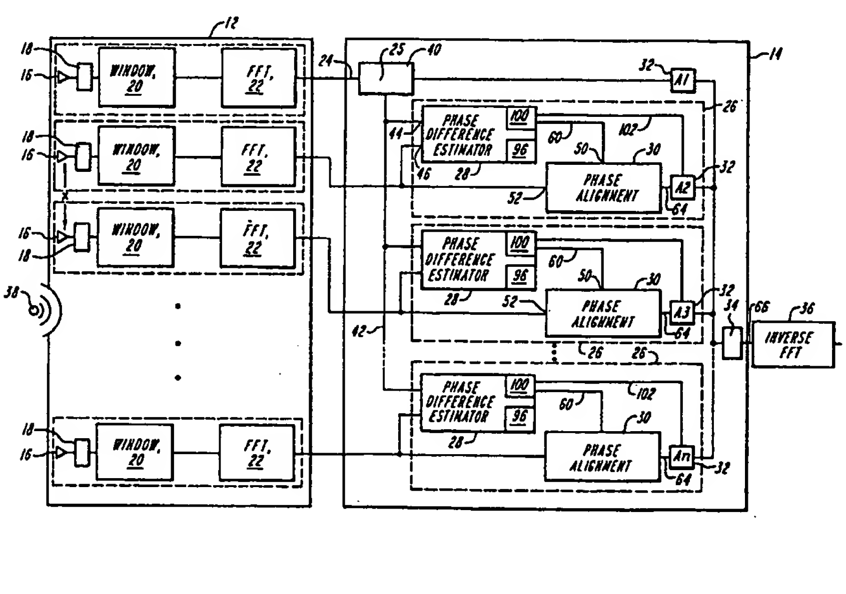
Assistant Examiner—Vivian W. Chang

Attorney, Agent, or Firm—Thomas J. Engellenner; Lahive & Cockfield

[57] **ABSTRACT**

Methods and systems for beamforming are disclosed that include a signal processor that can dynamically determine the relative time delays between a plurality of frequency-dependent signals. The signal processor can adaptively generate a beam signal by aligning the plural frequency-dependent signals according to the relative time delays between the signals. The signal processor can store one frequency-dependent signal as a reference signal and can align the remaining frequency-dependent signals relative to this reference signal. One advantage of the signal processor is that it can align the plural frequency-dependent signals generated by an array of microphones that can be arranged in a linear, two dimensional or three dimensional array and located in a room environment.

25 Claims, 6 Drawing Sheets



OTHER PUBLICATIONS

Silverman et al., "A Two–Stage Algorithm for Determining Talker Location From Linear Microphone Array Data," Computer Speech And Language, vol. 6, pp. 129–152 (1992).

Piersol, "Time Delay Estimation Using Phase Data," IEEE Transactions On Acoustics, Speech, And Signal Processing, vol. ASSP–29, No. 3, pp. 471–477 (1981).

Chan et al., "The Least Squares Estimation on Time Delay and Its Use in Signal Detection," IEEE Transactions On Acoustics, Speech, And Signal Processing, vol. ASSP–26, No. 3, pp. 217–222 (1978).

IEEE, International Conference On Acoustics, Speech And Signal Processing, p. 320, (1976).

IEEE International Conference On Acoustics, Speech And Signal Processing, vols. 3 of 4, p. 1699 (1987).

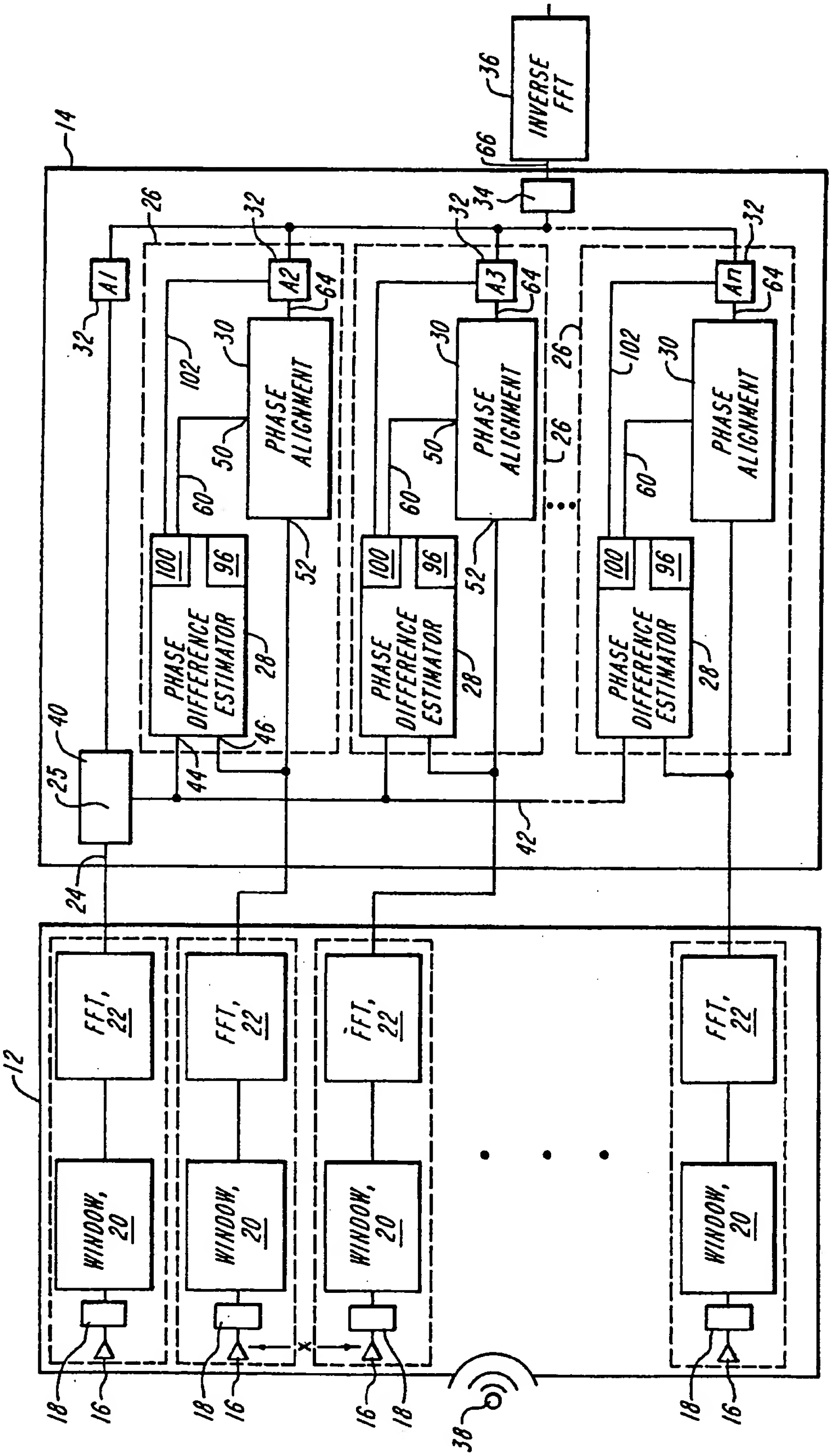

INVERSE FFT
PHASE ALIGNMENT
PHASE DIFFERENCE ESTIMATOR
FFT, 22
WINDOW, 20
A1
A2
A3
An
100
96
14
12
36
66
34
32
26
64
30
102
50
60
52
28
46
44
40
25
42
24
18
16
38
x

FIG. 1

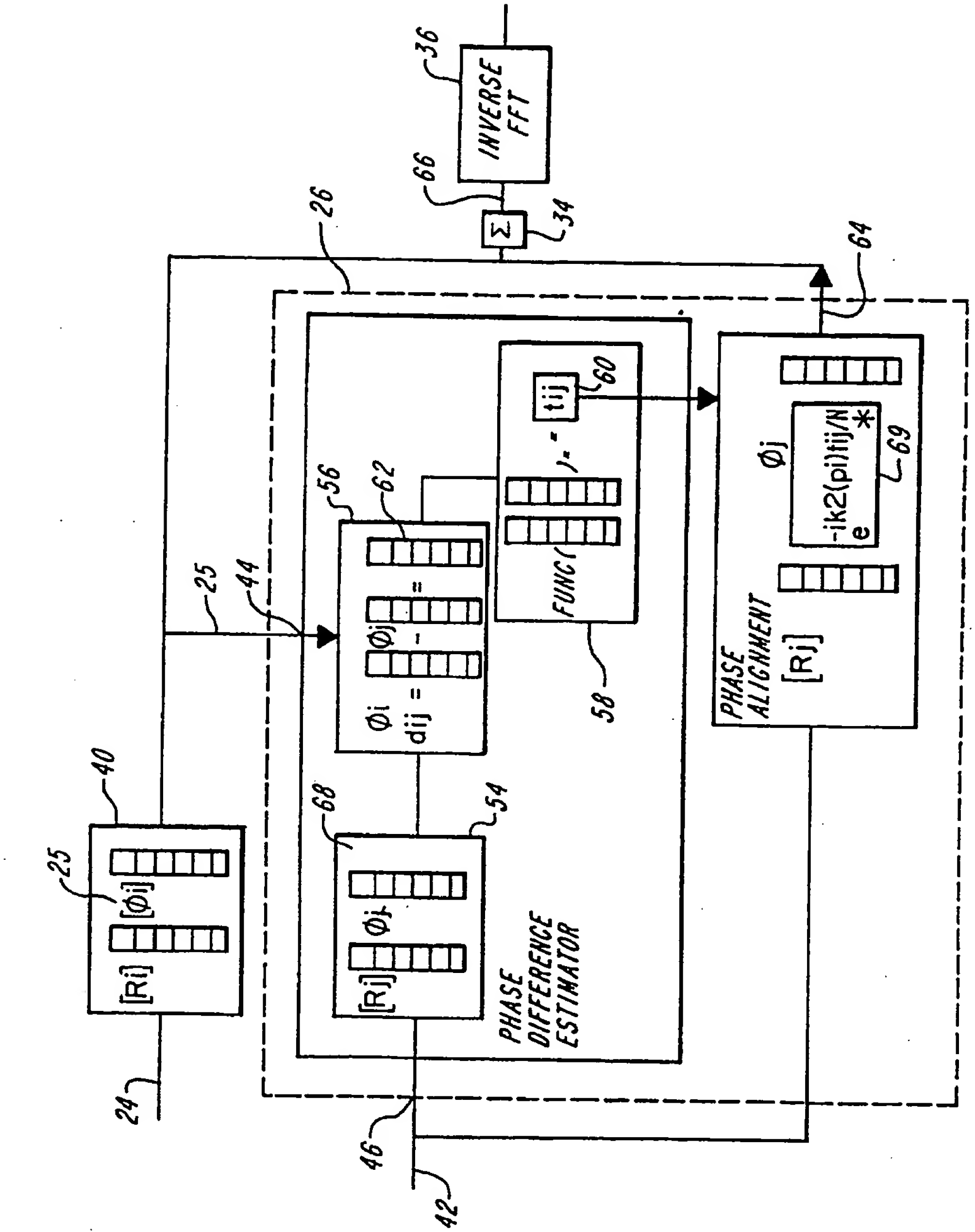

INVERSE FFT
PHASE DIFFERENCE ESTIMATOR
PHASE ALIGNMENT
FUNC(
[Ri]
[Rj]
[φi]
dij =
tij
Σ
24
25
26
34
36
40
42
44
46
54
56
58
60
62
64
66
68
69

FIG. 2

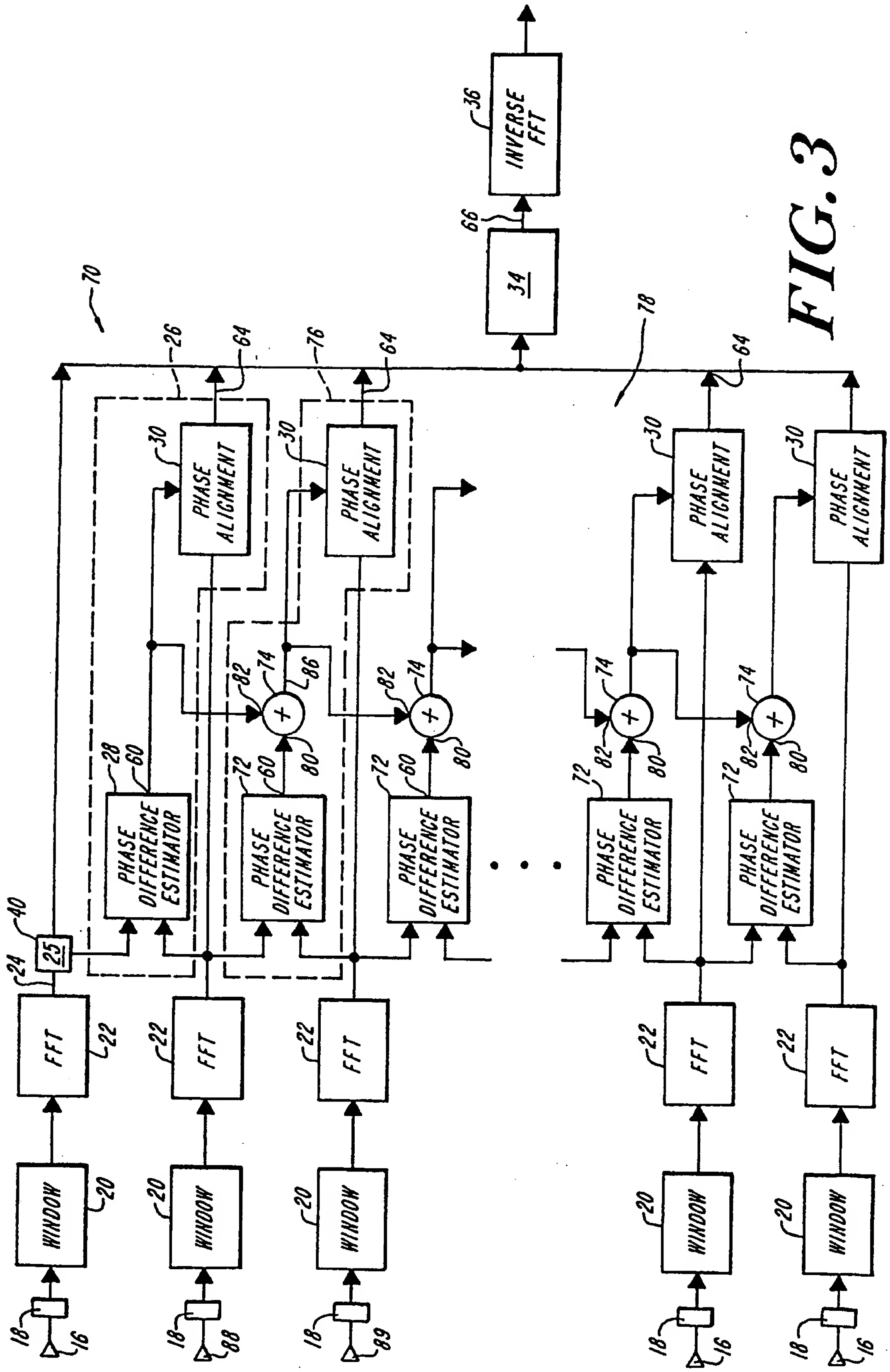

INVERSE FFT
34
PHASE ALIGNMENT
PHASE DIFFERENCE ESTIMATOR
FFT
WINDOW
25
+

FIG. 3

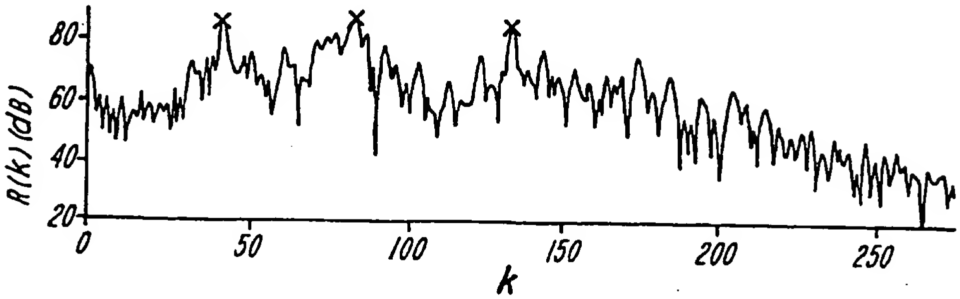


FIG. 4A

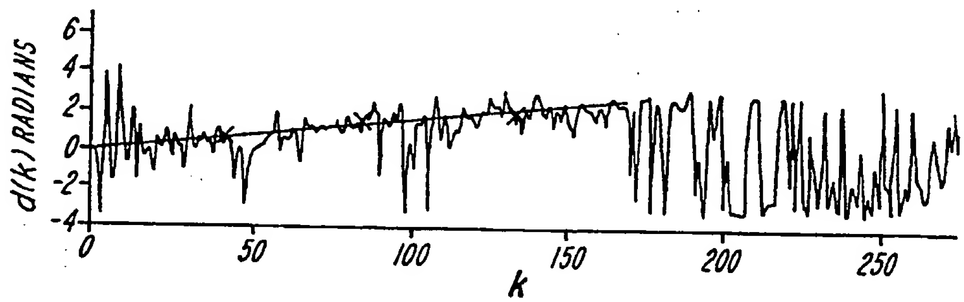


FIG. 4B

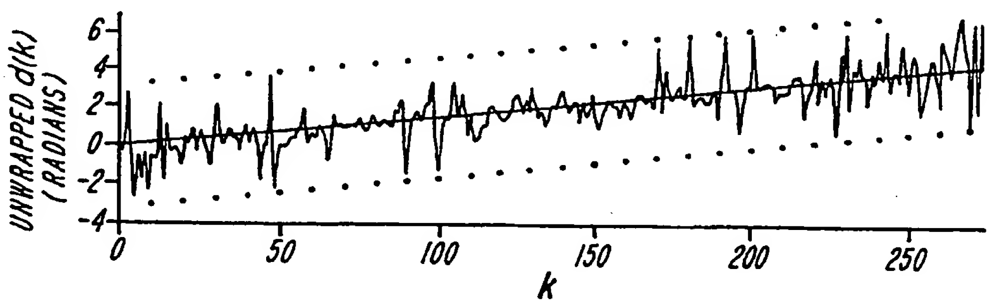


FIG. 4C

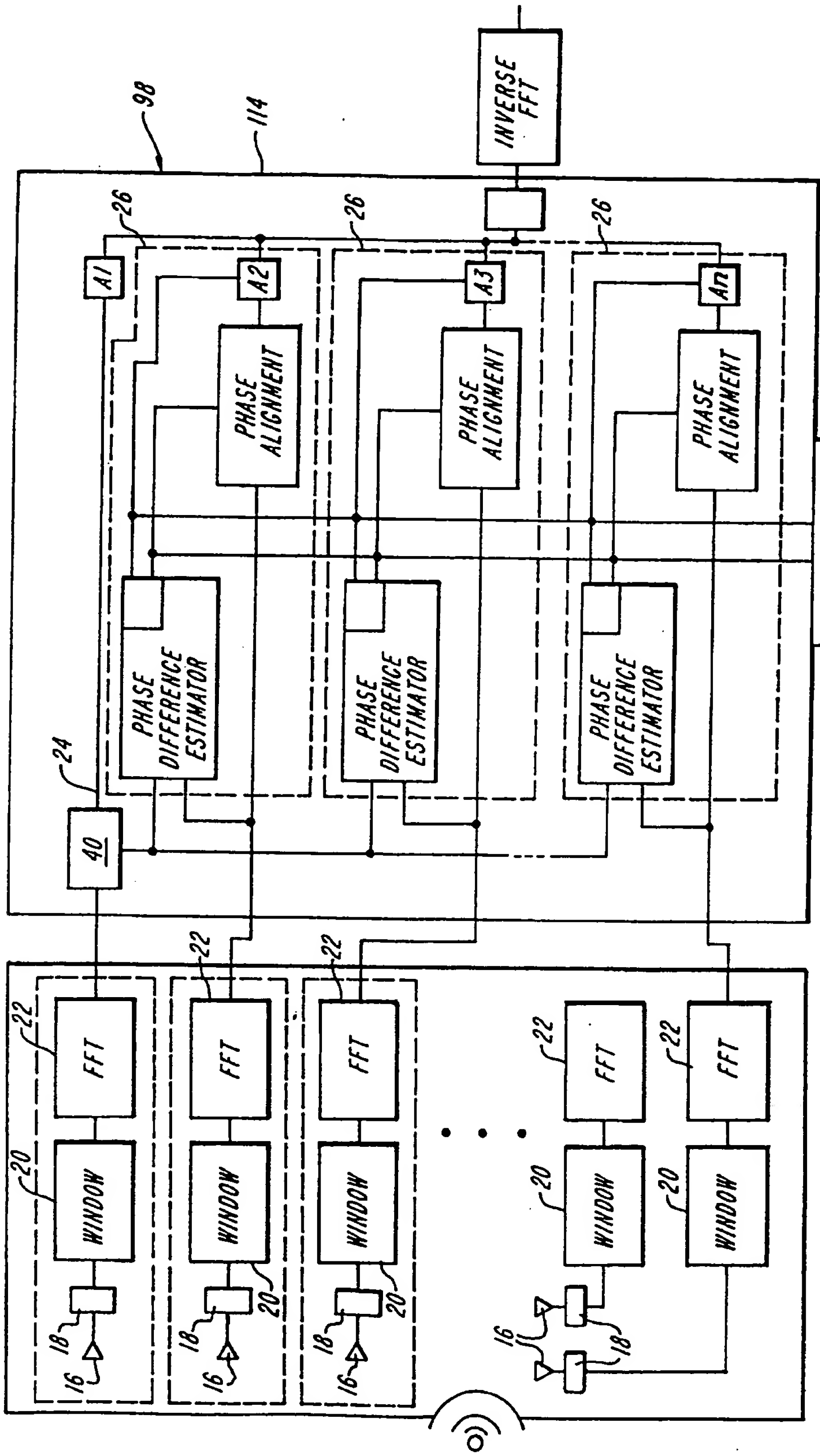

INVERSE FFT
98
114
26
A1
A2
A3
An
PHASE ALIGNMENT
PHASE DIFFERENCE ESTIMATOR
SOURCE LOCATOR
120
116
24
40
22
FFT
20
WINDOW
18
16

FIG. 5

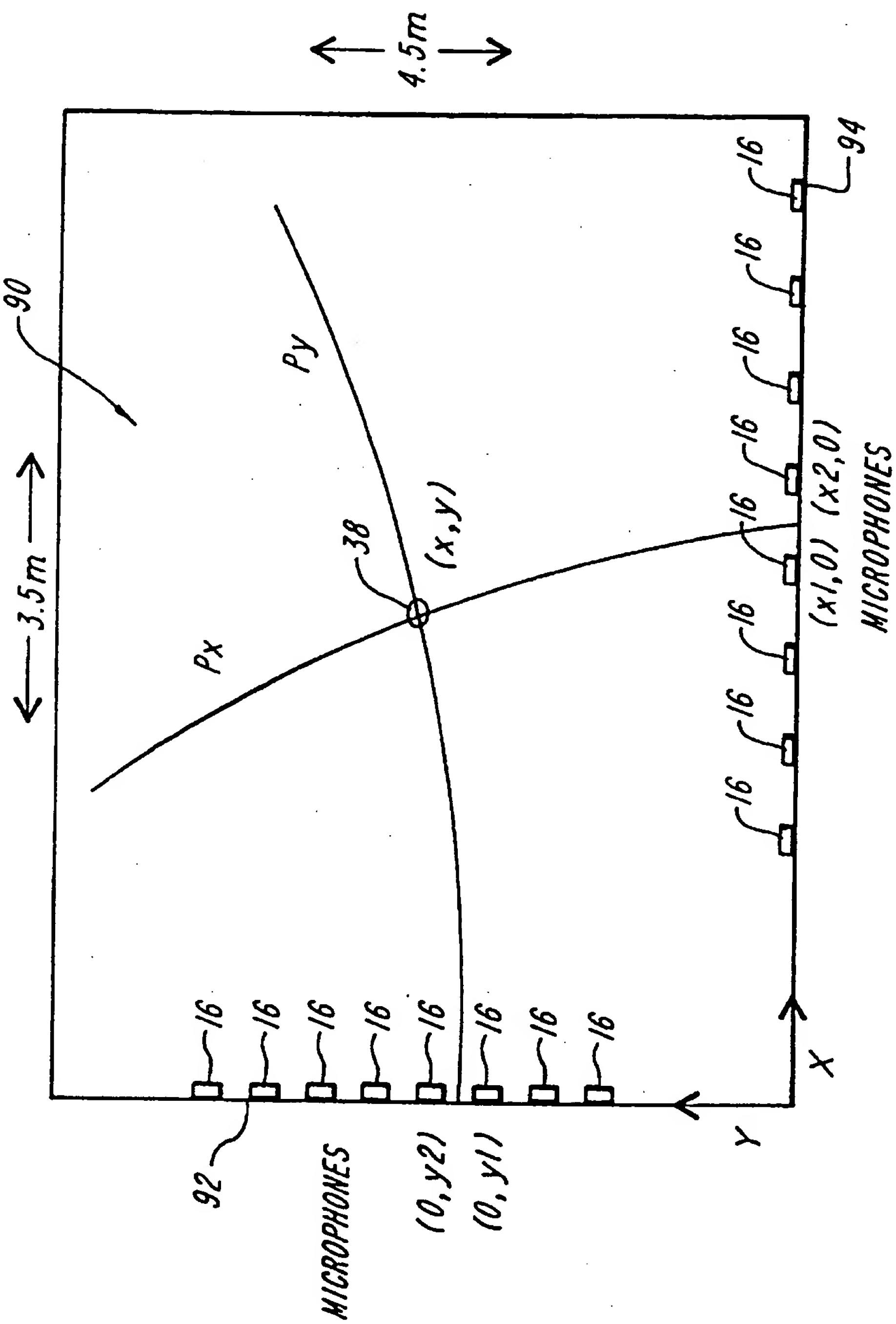

4.5m
3.5m
90
Py
Px
38
(x,y)
16
94
92
(x1,0) (x2,0)
MICROPHONES
(0,y2)
(0,y1)
MICROPHONES
X
Y

FIG. 6

METHODS AND APPARATUS FOR ADAPTIVE BEAMFORMING

This invention was made with government support under Grant/Contract No. MIP-9120843 awarded by the National Science Foundation. The government has certain rights in this invention.

FIELD OF THE INVENTION

The present invention relates to methods and apparatus for adaptive signal processing and, more particularly, to methods and apparatus for adaptively combining a plurality of signals, e.g., electrically represented audio signals, to form a beam signal.

BACKGROUND OF THE INVENTION

Many communication systems, such as radar systems, sonar systems and microphone arrays, use beamforming to enhance the reception of signals. In contrast to conventional communication systems that do not discriminate between signals based on the position of the signal source, beamforming systems are characterized by the capability of enhancing the reception of signals generated from sources at specific locations relative to the system.

Generally, beamforming systems include an array of spatially distributed sensor elements, such as antennas, sonar phones or microphones, and a data processing system for combining signals detected by the array. The data processor combines the signals to enhance the reception of signals from sources located at select locations relative to the sensor elements. Essentially, the data processor "aims" the sensor array in the direction of the signal source. For example, a linear microphone array uses two or more microphones to pick up the voice of a talker. Because one microphone is closer to the talker than the other microphone, there is a slight time delay between the two microphones. The data processor adds a time delay to the nearest microphone to coordinate these two microphones. By compensating for this time delay, the beamforming system enhances the reception of signals from the direction of the talker, and essentially aims the microphones at the talker.

A major factor in the effectiveness of these beamforming systems is the accuracy of the time delays necessary for aiming the sensor array. One known technique for determining the time delays necessary for aiming the sensor array employs a priori knowledge of the source position, the source orientation and the radiation pattern of the signal. Essentially, the data processor determines from the position of the source and, from the position of the sensor elements, a delay factor for each of the sensor elements. The data processor then applies such delay factors to the respective sensor elements to aim the sensor array in the direction of the signal source.

Although these systems work well if the position of the signal source is precisely known, the effectiveness of these systems drops off dramatically with slight errors in the estimated a priori information. For instance, in some systems with source-location schemes, it has been shown that the data processor must know the location of the source within a few centimeters to enhance the reception of signals. Therefore, these systems require precise knowledge of the position of the source, and precise knowledge of the position of the sensors. As a consequence, these systems require both that the sensor elements in the array have a known and static spatial distribution and that the signal source remains stationary relative to the sensor array. Furthermore, these beamforming systems require a first step for determining the talker position and a second step for aiming the sensor array based on the expected position of the talker.

Other techniques for determining the direction for aiming the sensor array rely on a priori information regarding the signal waveform and the signal radiation pattern. For example, radar systems use beamforming to transmit signals in a select direction. If an object is present in that direction, the signal reflects off the object and travels back toward the radar system. Therefore, the radar system is transmitting and receiving very similar signals. Furthermore, the data processor assumes that the objects are sufficiently distant from the sensor array that the incoming signals have a particular radiation pattern. The assumed radiation pattern can be a particularly simple pattern that reduces the complexity of the time delay computation.

The radar system capitalizes on the similarity of the transmitted and received signals by using signals that have features which facilitate signal processing. The data processor can directly compare the features of the received signal against the features of the transmitted signal and determine differences between the two signals that relate to the relative time delays between each sensor. Furthermore, the radar system can use the assumptions regarding the radiation pattern of the incoming signals to simplify the signal processing techniques necessary to calculate the time delays. The data processor then compensates for the respective time delays between each sensor element to aim the sensor array in the direction of the object.

Although these systems work well if the signal waveform is known, these systems less effective where the a priori information regarding the signal waveform is unavailable or insufficient to allow the received signals to be compared against a known signal waveform. Therefore, these systems are generally limited to active systems that both transmit and receive signals. Furthermore, these systems are less effective when assumptions regarding the radiation pattern cannot be made. Therefore, these systems are usually limited to those applications where the signal source is sufficiently distant from the sensor array that a signal pattern can be assumed.

A known technique for determining the direction of incoming signals without a priori information employs correlation strategies that compare signals received by the array at spatially distinct sensors to estimate the time delays between the sensors. The time delay information, along with assumptions about the radiation pattern, are used to estimate the location of the signal source. One example of correlation strategies for locating talker position with a microphone array in a near-field environment is set forth in Silverman et al., *A Two-Stage Algorithm for Determining Talker Location from Linear Microphone Array Data*, Computer Speech and Language, at 129–152 (1992). In general, the cross-correlation function of two signals received at two distinct sensors is computed and filtered in some optimal sense. The data processor includes a peak detector that detects the maximum value of the filtered signal. While the filtering criteria and the methods used for peak detection may vary considerably, these techniques are all based on maximizing the correlation between two received signals and determining from the detected peak the relative time delays between the associated sensors. Once the time delays are determined, techniques, such as triangulation, can be used to determine the location of the signal source.

Although these systems can work well, there is generally a trade-off between the accuracy of the time delay estimate

and the computational expense incurred by the procedure. Furthermore, there can be a tradeoff between the accuracy of the delay estimate and the rate at which the system can acquire the incoming signals. The cross-correlation function is a computationally intensive operation, and the accuracy of the peak data increases with the number of comparisons made during the correlation. In order to achieve a peak that is sufficiently accurate and well defined to identify precisely the position of the source, the computational burden can be prohibitive. Therefore, these systems can fail to produce the desired accuracy and update rate required for effective beamforming in a real-time environment.

In view of the foregoing, an object of the present invention is to provide improved signal processing methods and systems for combining a plurality of signals, and more particularly, to provide improved systems and methods for beamforming that dynamically determine the time delay estimates for a sensor array as part of the beamforming process.

A further object of the present invention is to provide systems and methods for real-time beamforming without the need of a priori information about the position of the signal source or knowledge of the signal radiation pattern.

Another object of the present invention is to provide signal processing systems and methods for adaptively aiming an array of sensor elements at a moving signal source.

A yet further object of the present invention is to provide signal processing systems and methods that can dynamically compensate for a sensor array that has a non-uniform or unknown spatial distribution of sensors.

A still further object of the present invention is to provide systems and methods for real-time beamforming without the need of a priori information about the signal waveform.

Still another object of the present invention is to provide computationally efficient systems and methods to determine the relative time delays between the signals received by the sensor elements of a sensor array and employ these delay estimates for computationally efficient beamforming and source location.

These and other objects of the invention are evident in the sections that follow.

SUMMARY OF THE INVENTION

The aforementioned objects are obtained by the present invention which provides in one aspect an adaptive beamforming apparatus which operates to combine a plurality of frequency-dependent signals to enhance the reception of signals from a signal source located at a select location relative to the apparatus.

In one embodiment, the beamforming apparatus connects to an array of sensors, e.g. microphones, that can detect signals generated from a signal source, such as the voice of a talker. The sensors can be spatially distributed in a linear, a two-dimensional array or a three-dimensional array, with a uniform or non-uniform spacing between sensors. In a typical practice, the sensor array can be mounted on a wall or a podium and the talker is free to move relative to the sensor array. Each sensor detects the voice audio signals of the talker and generates electrical response signals that represent these audio signals. The adaptive beamforming apparatus provides a signal processor that can dynamically determine the relative time delay between each of the audio signals detected by the sensors. Further, the signal processor includes a phase alignment element that uses the time delays to align the frequency components of the audio signals. The signal processor has a summation element that adds together the aligned audio signals to increase the quality of the desired audio source while simultaneously attenuating sources having different delays relative to the sensor array. Because the relative time delays for a signal relate to the position of the signal source relative to the sensor array, the beamforming apparatus provides, in one aspect, a system that "aims" the sensor array at the talker to enhance the reception of signals generated at the location of the talker and to diminish the energy of signals generated at locations different from that of the desired talker's location.

A beamforming apparatus constructed according to the present invention can include a signal processor that determines the relative time delay between a plurality of frequency-dependent signals. The signal processor can store one frequency-dependent signal as a reference signal and can align the remaining frequency-dependent signals relative to this reference signal. The reference channel can include a memory for storing one of the frequency dependent signals as a reference signal having a user selected phase angle. The reference channel can connect to a plurality of alignment channels, where each alignment channel couples to a respective one of the frequency-dependent signals. The alignment channels can operate to adjust the phase angle of each of the frequency-dependent signals in order to align the signals relative to the reference signal. Each alignment channel can have a phase difference estimator that generates a delay signal which represents the time delay between the reference signal and the respective signal connected to the alignment channel. The alignment channel can also include a phase alignment element that generates an output signal as a function of the delay signal, which has a magnitude that represents the magnitude of the respective signal and a phase angle that is adjusted into a select phase relationship with the reference signal. The signal processor can further include a summation element that couples to the alignment channels and to the reference channel. The summation element can generate a beam signal by summing the output signals with the reference signal.

The adaptive beamforming apparatus can include an array of spatially distributed sensor elements for generating the plurality of frequency-dependent signals. The sensor elements can be any one of a number of different types of elements capable of detecting a signal. Examples of such sensor elements include antennas, microphones, sonar transducers and various other transducers capable of detecting a propagating signal and transmitting the signal to the signal processor.

The sensor elements are spatially distributed to form an array for detecting a signal. Each sensor in the array can generate a single signal that represents the signal detected at that sensor element as a function of time. The spatial distribution of sensor elements can be unknown or non-uniform. The invention can be practiced with a linear array, a two dimensional array, or a three dimensional array.

In one embodiment of the invention, the reference channel of the signal processor can connect to the phase difference estimator of each alignment channel. In this practice, the phase difference estimator includes a memory for storing the reference signal and for storing the respective frequency-dependent signal associated with the respective alignment channel. The phase difference estimator has a processing means to generate the delay signal as a function of the reference signal and the respective frequency-dependent signal.

In an alternative embodiment, the signal processor can include interconnected alignment channels that determine

the relative time delay between spatially adjacent sensors. In this practice, the phase difference estimator can include a memory for storing the respective frequency-dependent signal of the associated alignment channel and the respective frequency-dependent signal of the second alignment channel. The memory can further store the delay signal of the second alignment channel. The phase difference estimator can include a summing element that generates a delay signal as a function of the signal associated with the respective alignment channel and delay signal of the second alignment channel.

In an alternative embodiment of the invention the signal processor can include a weighting element, that can increase or decrease the magnitude component of selected output signals. The weighting element can be a weighted averaging element that can affect the magnitudes of the output as a function of the number of output signals summed together.

In a further alternative embodiment of the present invention, an error detector is associated with each of the delay estimators and determines from the delay signals and the frequency-dependent signals, an error signal that represents the accuracy of the delay signals. The error signal can be used by the weighted averaging element to determine which of the output signals has an associated error signal that is larger than a user-selected error parameter. The summation means can effect the weighting of that output signal responsive to the error signal, including deleting that output signal from the signal summation.

In another further embodiment of the invention, the delay estimator generates a delay signal that represents the time delay between a reference signal and a respective one of the frequency dependent signals, by measuring the difference between the phase angle components to the frequency-dependent signals. In one embodiment the delay estimator measures the difference in phase angles between the reference signal and the respective frequency-dependent signal of that alignment channel. The delay estimator can calculate from the differences in phase angles and from the frequency associated with each phase angles, the relative phase shift between the two signals. In one embodiment of the invention, the delay estimator can further include a weighting system that multiplies the difference in phase angles of each frequency component of two respective signals, by the magnitude of that frequency component.

BRIEF DESCRIPTION OF THE DRAWINGS

The foregoing and other aspects of the invention may be more fully understood from the following description, when read together with the accompanying drawings in which like reference number indicate like parts in the several figures, and in which:

FIG. 1 illustrates a schematic block diagram of one embodiment of a beamforming apparatus constructed according to the present invention;

FIG. 2 illustrates a schematic block diagram of one alignment channel of the beamforming apparatus depicted in FIG. 1;

FIG. 3 illustrates an alternative embodiment of a beamforming apparatus constructed according to the present invention that includes phase difference estimators connected between spatially adjacent sensor elements;

FIG. 4 illustrates the operation of a delay estimator that includes an unwrapping element for limiting spatial aliasing;

FIG. 5 illustrates a further embodiment of the present invention that includes an orthogonal array of sensor elements;

FIG. 6 illustrates in more detail the orthogonal array of FIG. 1.

DETAILED DESCRIPTION

FIG. 1 depicts an adaptive beamforming apparatus **10** constructed in accord with the invention. The illustrated apparatus **10** includes a sensor array **12** and a signal processor **14**. The sensor array **12** includes the sensors **16**, sampling units **18**, window filters **20** and time-to-frequency transform elements **22**. The signal processor **14** includes a reference channel **24** and plural alignment channels **26**. Each alignment channel **26** includes a phase difference estimator **28**, phase alignment element **30** and an optional weighting element **32**. The illustrated system **10** further includes a summation element **34** and a frequency-to-time transform element **36**.

The illustrated sensor array **12** includes a plurality of sensor elements **16**. The sensors **16**, in the depicted embodiment, are arranged to form a spatially distributed linear array of sensors **16** each spaced apart by a distance X and arranged to receive input signals having signal components from a signal source, such as the target source **38**. In the illustrated embodiment, each sensor **16** is the front end of an reception channel that includes a sampling unit **18**, a window filter **20** and a time-to-frequency transform element **22** all connected in electrical circuit. Each of the illustrated reception channels is a distinct subsystem of the sensor array **12** and can operate simultaneously with and independently from the other reception channels.

Each sensor **16** detects signals, including signals generated from the target source **38**, and generates an electrical response signal that includes a component that represents the signal generated frown the signal source **38**. The sensors **16** in the sensor array **12** can be microphones, antennas, sonar phones or any other sensor capable of detecting a signal propagating from the source **38** and generating an electrical response signal that represents the detected signal.

Each illustrated sampling element **18** is in electrical circuit with one sensor **16** and generates a digital response signal by sampling the electrical response signal generated by the associated sensor **16**. The sampling element **18** can be a conventional analog-to-digital converter circuit of the type commonly used to sample analog electrical signals and generate digital electrical signals that represent the sampled signal. The sampling element **18** generates samples of the electrical response signal at a rate, f_{rate}, selected according to the application of the beamforming apparatus **10**. The sampling rate is generally determined according to the highest frequency component of the propagating signal of interest and according to the Nyquist rate. The sampling elements **18** are discussed in further detail below.

The window filter **20** can be a conventional digital window filter for selecting a discrete portion of a digital response signal. In the illustrated embodiment the window filter **20** is in electrical circuit with the output of the sampling element **18**, and generates a finite length digital signal by truncating the digital signal generated by the sampling unit **18**. In one embodiment, the window filter **20** can be a rectangular window filter that truncates the digital signal to a user-selected number of samples to represent the input signal detected by sensor **16**. Each discrete portion of the sampled signal is a frame of data that corresponds to the signal detected by the sensor **16** during a time period determined by the sampling rate and the number of samples present in the frame. The window filter **20** is discussed in further detail below.

In the depicted apparatus **10**, the window filters **20** are in electrical circuit with the time-to-frequency transform elements **22**. Each time-to-frequency transform element **22** can receive the data frames generated by filter **20** and transform each data frame into a frequency-dependent signal that represents the spectral content of the signal detected by the associated sensor **16** during the time period of the corresponding data frame. Each frequency-dependent signal can include a magnitude component, |R|, and a phase angle component, ϕ, for each frequency, ωn, in the spectral content of the transformed data frame. In one embodiment of the present invention, the frequency-dependent signals are stored in the apparatus **10** as complex arrays. Each complex array can include a storage cell that corresponds to a predetermined frequency, ω_n, and therefore can store the spectral contents of a data frame by filling the appropriate cell with the magnitude and phase angle of the corresponding frequency component in the spectral content of the data frame. For example:

|R|	ϕ
wo	wo
w1	w1
w2	w2
w3	w3
w4	w4
wn	wn

can be a complex array that represents the spectral content of one frame of data, and has a first array, |R|, that represents the magnitude component of each frequency, ω_n, and has a second array, ϕ, that represents the phase angle component of each frequency, ω_n. Other methods of storing or representing frequency-dependent signals should be apparent to one of ordinary skill in the art of signal processing and do not depart for the scope of the invention.

Therefore, the sensor array **12** generates from the target source **38** a plurality of frequency dependent signals, wherein each frequency-dependent signal is associated with one sensor **16**, and represents the signal generated by target source **38**, as "heard", by the associated sensor **16**. The time-to-frequency transform element **22** can be any of the commonly known signal processing techniques for efficiently computing the discrete fourier transform of a time domain signal. In a preferred embodiment of the invention the time-to-frequency transform element **22** is a Fast Fourier Transform element that performs the discrete fourier transform on the window input signal generated by filter **20**. It should be apparent to anyone of ordinary skill in the art of signal processing, that any efficient algorithm for transforming the input signal from the time domain to the frequency-domain can be practiced with the illustrated system, without the parting from the scope of the present invention.

The signal processor **14**, constructed according to the invention, combines the input signals detected by the sensor array **12** and essentially "aims" the sensor array **12** at a signal source, e.g. source **38**. The processor **14** "aims" the array **12** by generating a beam signal **66** that represents a combination of phase aligned input signals. The beam signal **66** enhances, i.e. increases the signal-to-noise ratio, of signals generated from a source at the position of target source **38** relative to the sensor array **12**.

The signal processor **14** has a reference channel **24**, plural alignment channels **26** and a summation element **34**. The reference channel **24** connects to one input channel and stores the frequency-dependent signal associated with that input channel in the memory element **40** as a reference signal **25**. The phase angle components of the reference signal can be defined as in-phase relative to the phase angle components of the other frequency-dependent signals. Each alignment channel **26** generates an output signal **64** representing the signal received at the associated sensor **16** phase aligned relative to the reference signal **25**. The phase aligned signals are combined to form the beam signal **66**.

The illustrated signal processor **14** is in electrical circuit with the sensor array **12** and receives the frequency-dependent signals generated by the time-to-frequency elements **22**. The signal processor **14**, depicted in FIG. **1**, is represented as circuit assemblies connected in electrical circuit. It should be apparent to one of ordinary skill in the art of signal processing that each circuit assembly depicted in FIG. **1** can be implemented as a software module and that the software modules can be similarly interconnected in a computer program to implement the signal processor **14** as an application program running on a conventional digital computer.

The illustrated signal processor **14** includes a plurality of channels each connected to a respective one of the frequency-dependent signals. In the illustrated embodiment, the signal processor **14** includes a reference channel **24** and a plurality of alignment channels **26**. The reference channel **24** has a storage element **40** for storing the reference signal **25** that represent the input signal detected by one of the sensors **16**. The memory **40** can store the reference signal **25** as a complex array. The storage element **40** is in electrical circuit via the conducting element **42** to each of the alignment channels **26**. The conducting element **42** connects to each of the phase difference estimators **28** in the alignment channels **26**. The phase difference estimator **28** of each of the alignment channels **26** has a second input **46** that is in electrical circuit with the output of a time-to-frequency transform element **22**.

With reference to FIG. **1** it can be seen that the alignment channels **26** of the illustrated signal processor **14** each connect to one time-to-frequency transform element **22**. The phase difference estimator **28** of each alignment channel **26** generates a delay signal **60** which approximates the time delay between the signal **25** detected by the sensor **16** associated with the reference channel **24** and the signal detected by the sensor **16** associated with alignment channel of the phase difference estimator **28**. This estimated delay signal **60** can be generated by any of the conventional time delay estimation techniques. These techniques can include cross-correlation algorithms with peak picking or frequency based delay estimators, including one preferred frequency based delay estimator that will described in greater hereinafter. For those delay estimators that include correlation techniques that operate in the time-domain, the phase difference estimator can include a frequency-to-time transform element to convert the magnitude and phase angle data of a data frame into a time dependent signal. A frequency-to-time transform element suitable for practice with the present innovation will be explained in greater detail herein after. However, any conventional domain transform algorithm or system can be practiced with the present invention without departing form the scope of the invention and such domain transform elements are considered within the ken of one of ordinary skill in the art of signal processing.

As further depicted by FIG. **1**, each alignment channel **26** of the signal processor **14** includes a phase alignment element **30** that connects in electrical circuit via the conducting element **48** to the output of the phase difference estimator **28**. The conducting element **48** carries the delay

signal **60** to the first input **50** of phase alignment element **30**. A second input **52** of phase alignment element **30** connects to the respective frequency-dependent signal of the respective input channel. As will be explained in greater detail hereinafter, the phase alignment element **30** can generate an output signal that is phase-aligned to the reference signal **25** stored in storage element **40**.

The output signals **64** of the depicted signal processor **14** are applied to optional weighting elements **32**. The weighting element **32** can increase or decrease the magnitude of the output signal. Each of the weighting elements **32** generate a weighted output signal that connects to the summation element **34**. The summation element **34** can sum together the weighted and phased aligned signals of each alignment channel and the weighted reference signal **25** of the reference channel **24**. The summation element **34** generates a beam signal **66**. The beam signal **66** represents a combination of phase aligned input signals that enhances, i.e. increases the gain, of signals generated from a source at the position of target source **38** relative to the sensor array **12**.

With reference to FIG. **2**, the construction and operation of a signal processor **14** constructed according to the embodiment shown in FIG. **1** can be described. FIG. **2** illustrates the reference channel **24**, the memory element **40**, a phase alignment channel **26**, that includes a phase difference estimator **28** and a phase alignment element **30**. The phase alignment element **30** and the memory element **40** are in electrical circuit to the summing element **34** that generates a signal transmitted over a conducting wire to the frequency-to-time transform element **36**. In the illustrated embodiment, the alignment channel **26**, including the phase difference estimator **28** and the phase alignment element **30**, aligns the frequency-dependent signal **68**, transmitted via conducting element **42**, to the reference signal **25**, stored in the data memory **40**.

In a first step, the phase difference estimator **28**, generates the delay signal **60** that represents the time delay between the reference signal **25** and the frequency-dependent signal **68**. In a second step, the phase alignment element **30**, calculates, for each frequency component of the frequency-dependent signal **68**, the phase shift:

k2(pi)tij/N

for that frequency component caused by the time delay. The phase alignment element **30** can align each frequency component of signal **68** as a function of the delay signal **60**, t_{ij}, and the frequency, 2(pi)k/N, where N can be the FFT size, and k can represent the frequency component, via the addition of the corresponding shift as given in the formula above, to the phase angle of the frequency-dependent signal **68**. The phase alignment element **30** generates the output signal **64**, that is aligned to the reference signal **25**, and that can be represented as a complex array, including a magnitude component and a phase angle component. In a final step, the aligned signal **68** and the reference signal **25** are combined by the summing element **34** to generate the beam signal **66**.

The phase difference estimator **28** illustrated in FIG. **2** includes the data memory **54**, a phase angle subtractor **56** and a delay estimator **58**. The illustrated phase difference estimator **28** is a frequency-domain phase difference estimator that generates the delay signal **60** that represents the relative time delay between the reference signal **25** stored in data memory **40** and the signal **68** stored the data memory **54**. The illustrated data memory **54** provides storage for a complex array having a magnitude component RJ and phase angle component ΦJ. The data memory **54** is in electrical circuit with the phase angle subtractor **56** that includes a data memory for storing the phase angle component, ΦI, of the reference signal **25** and for storing the phase angle component, ΦJ, of the signal **68** stored in data memory **54**. The phase angle subtractor **56** generates a signal **62** that represents the differences between the phase angles of the reference signal **25** and the phase angles of the respective frequency-dependent signal associated with that alignment channel **26**. The signal **62** can represent the phase angle difference as an array that has cells indexed by frequency. The difference signal **62** can be transmitted over a conducting element to the delay estimator **58**. In the illustrated embodiment the delay estimator **58**, which will be explained in greater detail hereinafter, generates the delay signal **60** as a function of the phase angle difference signal **62**.

The delay signal **60** connects via a conducting element to the phase alignment element **30**. As illustrated by FIG. **2**, the phase alignment element **30** is in electrical circuit with conducting element **42** to receive the frequency-dependent signal **68** associated with the alignment channel **26**. The phase alignment element **30** can include a phase shift element **69** that can generate a shift signal representative of the phase shifts for each of the frequency components of the signal **68**. The phase alignment element **30** can increment the phase angle ΦJ of the associated frequency-dependent signal by the shift signal. In one embodiment of the present invention, the phase alignment element **30** can be a programmable arithmetic-logic-unit that multiplies the phase angle of the associated frequency-dependent signal with the corresponding phase shift signal. However, it should be obvious to one of ordinary skill in the art of signal processing that the phase alignment element **30** can be implemented as a software module that includes programming structure for multiplying the phase angles of the signal **68** by the corresponding phase shift signals.

As further illustrated by FIG. **2**, the output signal **64** is transmitted via a conducting element to the summation element **34** along with the reference signal **25** stored in data memory **40**. The summation element **34** generates a beam signal **66** that represents the summation of the aligned output signals **64** from each of the alignment channels **26** in the signal processor **14** and the reference signal **25** stored in data memory **40**. The illustrated signal processor of FIG. **2** includes an optional frequency-to-time transform **36** element that generates a time-dependent signal that represents the beam signal **66**. In the illustrated embodiment the frequency-to-time domain transform element **36** is a inverse FFT of the type conventionally used to transform discrete signals from the time-domain to the frequency-domain.

With reference to FIG. **3**, one preferred embodiment of the present invention can be described. FIG. **3** depicts a beamforming apparatus **70** connected to a sensor array **12** and a signal processor **78**. The signal processor **78** includes a reference channel **24** that provides a data storage element **40** for storing one frequency-dependent signal associated with one of the sensors **16** as a reference signal **25** that includes a magnitude component and a phase angle component. The phase angle component of the reference signal **25** stored in the data memory **40** includes a phase angle corresponding to each one of the frequency components of the input signal detected by the sensor **16** associated with the reference channel **24**. The phase angles of the reference signal **25** can represent a reference phase for that frequency component of the signal generated by the source **38**. The storage element **40** generates an output signal that connects via a conducting element to the phase difference estimator **28** of the first alignment channel **26**. As can be seen with

reference to FIG. 3, the alignment channel **26** includes a phase difference estimator **28** and phase alignment element **30** constructed similarly to the previously described embodiment. The system **70** further includes a plurality of alignment channels **76** that include a phase difference estimator **72**, a summing element **74**, and a phase alignment element **30**. The alignment channels **76** connect between two input channels of the sensor array **12**. In the illustrated embodiment the alignment channels **76** preferably connect to spatially adjacent sensors in the sensor array **12**.

In the illustrated embodiment of FIG. **3**, the phase difference estimator **72** of each alignment channel **76** connects via conducting elements to the input channels of two spatially adjacent sensor elements to generate a delay signal **60** that represents the time delay between these two spatially adjacent sensors **16**. The alignment channel **76** further includes a summing element **74**. The summing element **74** has a first input **80** that connects via a conducting element to the output of the phase difference estimator **72**. The summing element **74** has a second input **82** that connects via a conducting element to the delay signal of a phase difference estimator associated with a sensor **16** that is spatially adjacent. The summing element **74** generates an output signal that is connected via a conducting element to the phase alignment element **30**.

As can be described with reference to FIG. **3**, the alignment channel **26** calculates the time delay between the reference signal **25** and the frequency-dependent signal generated by the spatially adjacent sensor **88**. A second alignment channel **76** calculates the time delay between the sensor **88** and the sensor **89**. The summing element **74** of the alignment channel **76** connects between the channel **26** and the channel **76** and can add together the two time delays to generate a cumulative delay signal **86**. The cumulative delay signal **86** represents the time delay between the sensor **16** of the reference channel **24** and the sensor **89** of the associated alignment channel **76**. As illustrated, each summing element **74** of each alignment channel **76** adds the cumulative delay signal **86** to the delay signal **60** generated by the phase difference estimator **72**. Therefore, the cumulative delay signal **86** references the each alignment channel **76** to the reference channel **24**.

The cumulative signal **86** generated by the summing element **74** represents the summed time delay between the reference signal **25** stored in data memory **40** and the frequency-dependent signal associated with the alignment channel **76**. The phase alignment **30** phase shifts the associated frequency-dependent signal by the total time delay represented by the signal **86** of summing element **74**. The phase shift added to each frequency component of the associated frequency-dependent signal aligns the associated frequency-dependent signal to the reference signal **25** stored in data memory **40**. The phase alignment element **30** generates an output signal **64** representative of the associated frequency-dependent signal phase aligned with the reference signal **25** stored in data memory **40**. The output signal of phase alignment element **30** is transmitted via a conducting element to the summing element **34**. As previously described, the summing element **34** sums the output signals generated by the alignment channels **26** and **76** with the reference signal stored in data memory **40**. The combined signals represents a beam signal **66** that can be transmitted by a conducting element to the optional frequency-to-time transform means **36**. The optional frequency-to-time transform element **36** can provide a output signal that represents the beam signal **66** as a time dependent signal.

The invention will now be further described with reference to one preferred embodiment that includes a frequency-domain delay estimator **58** and a linear array of microphones **16**. The frequency-domain delay estimator **58** aims the sensor array **12** by dynamically determining the time delay between two frequency-dependent signals to maximize the power in the beam signal **66** formed by the summation of the frequency-dependent signals. A signal processor **14** with this preferred frequency-domain delay estimator **58** is shown to be accurate over a wide range of signal-to-noise conditions and an effective basis for more complex acoustic-array applications, such as source detection and tracking procedures. Further, it is suitable for determining the time delay between wide-band frequency-dependent signals, where there is limited a priori knowledge of the spectral content of the signals.

The sensor array **12** includes a linear array of eight microphone sensors **16** distributed at 16.5 cm intervals along one wall of a room. The input signals detected by the microphones **16** are digitized simultaneously at 20 kHz by sampling units **18** of eight distinct input channels. The 20 kHz sampled input signals are windowed by window filter elements **20** into finite sequences. For each sequence the DFT is computed by the associated time-to-frequency transform element **22** and converted to a magnitude-phase representation. The choice of the window filter **20** and the size as well as the DFT length vary with the particular application and computational availability. One preferred window filter **20** is a 512-point Hanning window applied with zero padding for use with a 1024-point FFT as a time to frequency transform element **22**. The individual segments can be half-overlapping in time to facilitate reconstruction.

For each pair of spatially consecutive microphones **16**, the phase angle subtractor **56** calculates the phase angle differences between corresponding frequencies and generates the signal **62**, $d_{ij}(k)$. Each frequency component of the frequency-dependent signals can be represented by:

$$\Omega = \frac{2\pi k}{N}$$

where N is the DFT length, k=0, 1, . . . , N- 1, and Ω is angular frequency. R(k) represents the spectral magnitude component of the frequency dependent signal. A phase delay signal **60**, τ_{ij}, is then computed according to the function:

$$\tau_{ij} = \frac{\sum_{k=0}^{K} |R(k)|^2 \left(\frac{2\pi k}{N}\right) d_{ij}(k)}{\sum_{k=0}^{K} |R(k)|^2 \left(\frac{2\pi k}{N}\right)^2}$$

where R(k) can represent, in one embodiment, the geometric mean of the magnitude components of the frequency-dependent signals, R_i, R_j. It should be obvious to one of ordinary skill in the art of signal processing that values for R(k) can be computed in using other statistical techniques, including determining the median of plural signals, weighted averaging and other techniques that can improve signal-to-noise rejection and error estimate.

The frequency-domain delay estimator **28** can include an optional unwrapping element **96**. The unwrapping element **96** is understood to resolve any spatial aliasing in the delay signal **60**. In one embodiment the delay estimator **28** includes an unwrapping element **96** that can generates the delay signal **60**, τ_{ij}, in three iterations, each of which generates an increasingly accurate estimate of the time delay between the signals. The accuracy of the delay estimate is understood to depend upon the limits of summation in the above equation. In general, the delay estimate tends to

converge upon the true delay more precisely as the number of terms in the summation is increased. Therefore, it is preferred to sum over k=0, 1, . . . , k_{max} where k_{max} corresponds to the highest frequency of interest. For speech, a reasonable cutoff can be 5.4 kHz with

$$k_{max} = \left\lfloor \frac{5400 * N}{f_{rate}} \right\rfloor .$$

However, the 2 πm phase ambiguity in the delay signal **60** can restrict the region in which the phase angle difference signal **62** is understood to vary in a linear fashion and therefore limits the upperbound limit of the summation index. One preferred unwrapping element **96** generates a delay signal **60** by providing two initial estimates of the delay signal **60**.

The unwrapping element **96** can generate an initial estimate for the delay signal **60**, τ_{ij1}, by deterring a first frequency range over which spatial aliasing is understood not to occur. The first range, K, is determined by:

$$K = K_{ij1} = \min\left\{ \left\lfloor \frac{c * N}{2 * |m_j - m_i| * f_{rate}} \right\rfloor , k_{max} \right\}$$

where c is the propagation speed of the input signals, and $|m_j - m_i|$ represents the spatial distance between the microphones **16**. The minimum of the two solutions can be used for K.

The unwrapping element **96** can generate a second estimate of the delay signal **60** by computing the delay signal **60** over the range determined by:

$$K = K_{ij2} = \min\left\{ \left\lfloor \frac{N}{2(|\tau_{ij1}| + \epsilon} \right\rfloor , k_{max} \right\}$$

The error term, ε, can be included in the above expression to compensate for the inaccuracy of the initial estimate of the delay signal **60**, t_{ij1}. Nominal values for ε range from 0.5 to 2 samples, depending on the expected accuracy of the initial estimate.

In a third iteration, the unwrapping element **96** uses the second estimate of the delay signal **60**, τ_{ij2}, to unwrap the phase angle difference signal **62**, $d_{ij}(k)$, and then a final estimate for the delay signal **60** can be computed over the entire frequency range of interest ($K=K_{max}$). The phase angle differences in signal **62** should vary linearly in frequency with variations in linearity due to additive noise in the sensor signal. The delay estimator **58** can examine the phase angle differences as a function of frequency, and given the second estimate of the delay signal **60**, unwrap the phase differences that evidence a 2πm phase ambiguity. It is preferred that the unwrapping depend upon an accurate estimate of π_{ij}, which is typically not available until the end of the second iteration.

The iterative procedure of the unwrapping element **96** is illustrated in FIG. **4**. The upper graph is a plot of spectral magnitudes in dB for the frequency-dependent signal, the middle graph displays the original phase angle difference signal **62** used for the first two iterations, and the bottom graph is the unwrapped phase angle difference signal **62** applied in the final iteration of the algorithm. In each case, the horizontal axis is the first 275 points of the DFT, corresponding to 0 through 5.4 kHz. In the initial stage, K_{ij1}=53, which when used as the upper bound of the summations for the initial estimate of delay signal **60**, and generates a time delay in samples of τ_{ij1}=1.513 samples. This estimate of the delay signal **60** is then used to calculate the range of summation for the second iteration. Using an error term ε=1.5 samples, K_{ij2}=169 and the second delay estimate for signal **60** is found to be τ_{ij2}=2.579 samples. The delay signal **60** may be viewed as the slope of the line that fits these points in a weighted mean squared sense. In the second graph, the phase wrapping ambiguity is apparent and the graph does not appear to be linear. In the third iteration, the phase differences in the signal **62** are unwrapped by the unwrapping element **96** and plotted on the lower graph. The unwrapping algorithm places each phase angle difference within π radians of the slope line by adding/subtracting integer multiples of 2π. The dotted lines in the lower graph represent the boundaries of the unwrapping algorithm. The final delay signal **60**, τ_{ij}, is then calculated with the unwrapped phase angle difference signal **62**, $d_{ij}(k)$, over the entire frequency range (k=0,1, . . . , k_{max}).

The frequency-domain delay estimator has several advantages over its time-domain counterpart. It is computationally simple, does not necessitate the use of search methods, and has precision independent of sampling rate.

With reference again to FIG. **1**, a further embodiment of the present invention, that includes an error detection element **100** can be described. The delay estimator **28** of FIG. **1** includes an optional error detection unit **100** that is in electrical circuit the weighting element **32**. The error detection unit **100** can generate an error signal **102** that represents the accuracy of the delay signal **60** generated by the phase difference estimator **28**. In one preferred embodiment of the invention, the weighting element **32** can affect the weighting of the aligned output signal **64** responsive to the error signal **102**. The weighting element **32** can include a user-selected error parameter. The weighting element **32** can compare the generated error signal **102** with the user-selected error parameter and generate a weighting parameter for the associated output signal **64** as a function of the error signal **102** and the user-selected error parameter.

In one preferred embodiment of the error detection unit **100**, the detection unit **100** includes a data processor that generates the error signal **102** as a function of the phase angle difference signal **62** and the magnitude components of the frequency dependent signal. In one example the error signal **102** is computed from:

$$\mathrm{Error}_{Norm}(\tau_{ij}) = \frac{\sum_{k=0}^{K} |R(k)|^2 \left(\left(\frac{2\pi k}{N} \right) \tau_{ij} - d_{ij}(k) \right)^2}{\sum_{k=0}^{K} |R(k)|^2}$$

The error signal **102** can provide a useful means for evaluating the significance of a delay signal **60**. A relatively large error signal **102** can indicate that the predicted delay signal **60** is inaccurate, as would be expected during times when there are no input signals in-coming to the sensor array **12**. A small value can demonstrate that the delay signal **62** is a good measure of the relative time delay between the sensors **16**.

In one embodiment, a normalized version of this error signal **102** can be calculated and compared to a user-selected parameter that represents an environmentally dependent threshold to determine if the delay signal **60** is valid. Environmentally dependent factors can include background noise, deviations between sensor performance and other similar factors.

In another preferred embodiment, the error detection unit **100** generates a signal that represents the geometric mean of the individual magnitudes of the frequency-dependent signals, $|R(k)| = \sqrt{|R_i(k)||R_j(k)|}$, and uses this mean to compute the error signal **102**. This preferred embodiment is under-

stood to be more resistant to noise and gain differences between the sensors **16**.

In a further embodiment of the present invention, depicted in FIG. **5**, a beamforming apparatus **98** according to the invention can be constructed having an orthogonal array **90** of sensor elements **16**. The beamforming apparatus **98** according to this embodiment of the invention determines the position of target source **38** through a series of triangulation calculations which require knowledge of the signal's relative delay when projecting onto a pair of microphone receivers.

The beamforming apparatus **98** can include the orthogonal array **90**, and a signal processor **114**. The orthogonal array **90** can include a plurality of sensor elements **16** each connected to an input channel that includes a sampling unit **18**, a window filter **20** and a time-to-frequency transform element **22**. The signal processor **114** can include a reference channel **24** and plural alignment channels **26**. Each alignment channel **26** includes a phase difference estimator **28**, phase alignment element **30** and an optional weighting element **32**. The signal processor **114** can further include a source locator unit **116**, in electrical circuit with each of the phase difference estimators **28**, a summation element **34** in electrical circuit with each of the phase alignment elements **30** and a frequency-to-time transform element **36** in electrical circuit with the summation element **34**. As will be explained in greater detail hereinafter, the source locator unit **116** generates an output signal **120** that represents the location of the detected source, e.g., source **38**, relative to the sensor array **90**.

FIG. **6** illustrates the orthogonal array **90** that includes sensor elements **16** distributed in two independent arrays including a horizontal array **94** and a vertical array **92**. An orthogonal array is preferred for its stability in evaluating both the x and y positions although other transverse array configurations can be practiced with the present invention. Further, it should be apparent to one of ordinary skill in the art of signal processing that the array **90** can include third array of sensors **16** disposed above or below the plane formed by the orthogonally arranged arrays **92** and **94**. The third array can configured into the system in the manner of arrays **92** and **94** and can yield time delay information, related to a third dimension, or coordinate of the source **38**, for example height.

While either linear array **92** or **94** may be used to evaluate both the x and y coordinates of the source position, the triangulation procedure is understood to be most effective if position coordinates are determined by the array in the direction normal to the source. For example, using only the sensors **16** in the array **94** is effective for evaluating the x-coordinate of the source location **38**, but not as accurate at finding the y-coordinate. By combining both axes in the triangulation procedure, the estimate is equally sensitive in either direction.

Each sensor **16** detects signals, including signals generated from the target source **38**, and generates an electrical response signal that includes a component that represents the signal generated from the signal source **38**. The sensors **16** sensor array **90** can be microphones, antennas, sonar phones or any other sensor capable of detecting a propagating signal and generating an electrical response signal that represents the detected signal.

The source locator **116** can generate the position signal **120** that represents the position of the source **38** relative to the sensor array **90**. In one preferred embodiment of the source locator **116**, at least four phase difference estimators **28** transmit delay signals **60** to the source locator **116**. Preferably the delay signals **60** transmitted to the source locator **116** represent the time delay between two spatially adjacent sensors **16** in array **94** and two spatially adjacent sensors **16** in array **92**. With reference to FIG. **6**, the generation of position signal **120** can be explained. Given four sensors **16**, one pair on the x-axis array **92** at positions x1 and x2 and another pair on the y-axis array **94** at y1 and y2, the curves Px and Py represent the loci of points px∈Px and py∈Py such that:

$$|px - x2| - |px - x1| = \delta_x$$
$$|py - y2| - |py - y1| = \delta_y$$

where δ_x and δ_y are constants such at $|\delta_x| \leqq |x2-x1|$ and $|\delta_y| \leqq |y2-y1|$. The curve Px can be interpreted as the set of locations which produce the same relative delay between x1 and x2. This relative delay, represented by the delay signal **60**, τ_x (in samples) can be related to δ_x by the following relation:

$$\delta_x = \frac{\tau_x * c}{f_{rate}}$$

Where f_{rate} is the sampling rate of the sampling elements **18**. Py and δ_y may be regarded similarly with respect to the sensors **16** on the y-axis array **94**.

The intersection of Px and Py represents a unique source location that produces relative delay signals **60**, τ_x and τ_y, between the respective sensor **16** pairs. The source locator unit **116** can generate the position signal **120** by estimating the relative delays at each sensor pair, and generating the curves Px and Py and find their intersection. Given that Px and Py represent one half of the hyperbolas, the intersection of Px and Py may be solved for algebraically. The simultaneous solution of the hyperbola equations reduces to finding the roots of a fourth order polynomial. From these four roots, the real root which corresponds to the actual coordinate pair (x,y) of the source location can be identified. This is can be accomplished by noting that the four intersection points of these two hyperbolas are each located in a distinct quadrant of the x-y plane. These four quadrants are demarcated by the lines y=(y1+y2)/2 and x=(x1+x2)/2. The proper quadrant may be chosen directly from the signs of the δ_x and δ_y terms.

In one preferred embodiment of the source locator **116**, the locator **116** can select which sensor pairs and delay signals **60** to use to generate the position signal **120**. For eight sensors **16** there are 28 subsets of two which corresponds to 28^2=784 combinations of the x-y axes sensor pairs. The first restriction imposed is to consider only pairs of sensors **16** that are spatially contiguous. The second constraint is to consider only those delay signals **60** with an associated normalized error less than a certain threshold. The error signal **102** of each error unit **100** can be transmitted by a conducting element to the source locator **116**. The source locator can compare the error signal **102** against a user-selected error parameter. If the comparison indicates a large error, then that indicates that the delay signal **60** is either inaccurate, the single source model does not apply, or this is a region of silence. In the first two cases the position signal **120** generated by the source locator **116** is a low quality estimate of the position. In the final case the position signal is meaningless as a position signal **120** but does indicate the presence of a signal source **38**.

In the preferred embodiment, the source locator **116** connects to each delay estimator **28** and, for each array **92** and **94**, collects the delay signals **60** and corresponding error signal **102** for each set of sensor pairs with less than a user-selected error-threshold. The source locator **116** orders each set by increasing normalized error as represented by the

error signal **102**. If either set is empty then no position signal **120** is generated. If either set has sensor pairs with error signals **102** below the user-selected error parameter, then the source locator **116** generates a position signal for a user-selected number of sensor pairs. The position signal **120** can be generated as the mean of several position estimates.

The source locator unit **116** can be a conventional electrical circuit card that includes arithmetic and logic circuits for generating from delay signals **60** of the phase difference estimators **28**, a position signal that represents the position of the source **38** relative to the sensor array **90**. The source locator unit **116** can also be a conventional data processor, such as a engineering workstation of the type sold by the SUN Corporation, having an application program for generating from the delay signals **60** of the phase difference estimators **28**, a position signal that represents the position of the source **38** relative to the sensor array **90**.

Described above are improved methods and apparatus for combining a plurality of signals to generate a beam signal for enhancing the reception of signals at a select position relative to an array of sensor elements. The invention has been described with reference to preferred, but optional, embodiments of the invention that achieve the objects of the invention set forth above.

Thus, for example, a steerable array of microphones has been described that has the potential to replace the traditional microphone as the input transducer system of speech data. An array of microphones has a number of advantages over a single-microphone system. It may be electronically aimed to provide a high-quality signal from a desired source location while it simultaneously attenuates interfering talkers and ambient noise. In this regard, an array has the ability to outperform a single, highly-directional microphone. An array system does not necessitate local placement of transducers, will not encumber the talker with a hand-held or head-mounted microphone, and does not require physical movement to alter its direction of reception. These features make it advantageous in settings involving multiple or moving sources. Furthermore, it is capable of activities that a single microphone cannot perform, namely the automatic detection, location, and tracking of active talkers in its reception region. Existing array systems have been used in a number of applications. These include teleconferencing, speech recognition, speech acquisition in an automobile environment, large-room recording-conferencing, and hearing aid devices. These systems also have the potential to be beneficial in several of other environments, the performing arts and sporting communities, for instance.

The above described embodiments have been set forth to describe more completely and concretely the present invention, and are not to be construed as limiting the invention. Thus, for example, the invention can be practiced as a radar system having two dimensional array of antenna elements disposed at non-uniform spacing in an plane. The array can couple to a signal processor constructed according to the present invention, that can align each of the signals received by the antenna relative to each other. Additionally, the radar system can include a source locator unit that determines from the relative time delays between the antennas the position of the source relative to the antenna array.

It is further intended that all matter and the description and drawings be interpreted as illustrative and not in a limiting sense. That is, while various embodiments of the invention have been described in detail, other alterations which will be apparent to those skilled in the art are intended to be embraced within the spirit and scope of the invention.

In view of the foregoing, what is claimed is:

1. Signal processing apparatus for combining a plurality of frequency-dependent signals wherein each frequency-dependent signal has a magnitude component and a phase angle component, said apparatus comprising

reference means for defining one of said frequency-dependent signals as a reference signal having a user-selected phase angle,

a plurality of alignment means, each coupled to a respective one of said frequency-dependent signals, for adjusting the phase angles of said signals relative to said reference signal, said alignment means having

phase difference estimator means for generating a delay signal representative of a time delay between said reference signal and said frequency-dependent signal, and

phase alignment means for generating, as a function of said delay signal, an output signal having a magnitude component representative of the magnitude component of said frequency-dependent signal and having a phase angle component adjusted to a select phase relationship with said reference signal, and

summation means, coupled to said plurality of alignment means for summing together said phase aligned output signals to generate a beam signal.

2. Apparatus according to claim 1 further comprising

means for generating said plurality of frequency-dependent signals, said means including

an array of spatially distributed sensor elements, wherein each sensor element includes means for detecting a signal and generating a respective one of said plural frequency-dependent signals to represent said signal detected at said spatially distributed sensor element.

3. Apparatus according to claim 2 wherein

said array includes a linear array of spatially distributed sensor elements.

4. Apparatus according to claim 2 wherein

said array includes a two-dimensional array of spatially distributed sensor elements.

5. Apparatus according to claim 2 wherein

said array includes a three-dimensional array of spatially distributed sensor elements.

6. Apparatus according to claim 1 wherein said phase difference estimator means includes

means for generating said delay signal as a function of said reference signal and said respective one of said frequency-dependent signal.

7. Apparatus according to claim 1 wherein said phase difference estimator means couples to a delay signal of a second alignment means and includes

summing means for summing said delay signals to generate a signal representative of the time delay between said respective one of said frequency-dependent signal and said reference signal.

8. A signal processing apparatus according to claim 1 further comprising

weighting means, connected to one or more phase alignment means, for increasing or decreasing the magnitude component of each of said output signals.

9. A signal processing apparatus according to claim 1 further comprising

weighted averaging means, connected to at least a portion of said phase alignment means, for increasing or decreasing the magnitude component of said output

signals as a function of a normalizing factor representative of the number of output signals summed together.

10. Signal processing apparatus for combining a plurality of frequency-dependent signals wherein each frequency-dependent signal has a magnitude component and a frequency component, said apparatus comprising

reference means for defining one of said frequency-dependent signals as a reference signal having a user-selected phase angle,

a plurality of alignment means, each coupled to a respective one of said frequency-dependent signals, for adjusting the phase angles of said frequency-dependent signals relative to said reference signal, said alignment means having

storage means for storing a magnitude component and a phase angle component of said frequency-dependent signal,

delay estimator means for generating, as a function of the difference in phase angles of two frequency-dependent signals, a delay signal representative of a time delay between said reference signal and said frequency-dependent signal, and

phase alignment means for generating as a function of said delay signal, an output signal having a magnitude component representative of the magnitude component of said frequency-dependent signal and having a phase angle adjusted to a select phase relationship with said reference signal, and

summation means, coupled to said plurality of alignment means and having means for summing frequency-dependent signals, for generating a beam signal representative of a summation of said output signals.

11. A signal processing apparatus according to claim **10** wherein said delay estimator includes weighting means for generating as a function of said magnitude components of said frequency-dependent signal, said difference in phase angles.

12. A signal processing apparatus according to claim **10** further including

error detection means for generating, as a function of said delay signal and said phase angle component of said frequency-dependent signal, an error signal representative of the accuracy of said delay signal.

13. A signal processing apparatus according to claim **12** wherein said summation means includes means for monitoring said error signal to adjust said beam signal responsive to an error signal larger than a user-selected error-parameter.

14. A signal processing apparatus according to claim **12** further comprising

means for generating said error signal as a function of the geometric mean of the magnitude components of two frequency-dependent signals.

15. A beamforming apparatus for combining a plurality of frequency-dependent signals wherein each frequency-dependent signal has a magnitude component and a phase angle component comprising

means for generating said plurality of frequency-dependent signals, having an array of spatially distributed sensor elements, wherein each sensor element includes transducer means for detecting a signal and for generating a respective one of said plural signals to represent said signal detected at said spatially distributed sensor element,

reference means for storing one of said frequency-dependent signals as a reference signal having a user-selected phase angle,

a plurality of alignment means, each coupled to a respective one of said frequency-dependent signals, for adjusting the phase angle components of said frequency-dependent signals relative to said reference signal, said alignment means having

storage means for storing said magnitude component and said phase angle component of said frequency-dependent signal,

delay estimator means for generating, as a function of the difference in phase angles of two frequency-dependent signals, a delay signal representative of a time delay between said reference signal and said frequency-dependent signal, and

phase alignment means for generating as a function of said delay signal, an output signal having a magnitude component representative of the magnitude component of said frequency-dependent signal and having a phase angle component adjusted to a select phase relationship with said reference signal, and

summation means, coupled to said plurality of alignment means and having means for summing frequency-dependent signals, for generating a beam signal representative of a combination of said output signals.

16. Apparatus according to claim **15** wherein

said array includes a linear array of spatially distributed sensor elements and said detection means includes means for detecting audio signals.

17. Apparatus according to claim **15** wherein

said array includes a linear array of spatially distributed microphones of the type amenable for detecting audio signals.

18. Apparatus according to claim **15** wherein

said array includes digital conversion means, coupled to each of said sensor elements, for generating said respective signal as digital electrical signal.

19. Apparatus according to claim **18** wherein

said array includes window filter means, coupled to each of said sensor elements, for generating said respective signal to represent a discrete portion of said digital electrical signal.

20. Apparatus according to claim **18** wherein

said array includes a 512 point hanning window filter means, coupled to each of said sensor elements, for generating said respective signal to represent a 512 point portion of said digital electrical signal.

21. Apparatus according to claim **15** wherein said array further comprises

time-to-frequency transform means, coupled to each of said sensor elements, for generating said respective signal as a frequency-dependent representation of said detected signal.

22. Apparatus according to claim **21** wherein said frequency transform means includes

fast fourier transform means for generating a plurality of fourier coefficients representative of at least a portion of the spectral content of said detected signal.

23. Apparatus according to claim **15** wherein said delay estimator further comprises

spatial aliasing filter means for generating said delay signal as a function of the spatial distribution of said sensor elements.

24. Apparatus according to claim **15** where in said summation means further comprises

frequency-to-time transform means, coupled to said signal summation means, for generating said beam signal as a time-dependent signal.

25. Apparatus according to claim 15 wherein

said array of spatially distributed sensor elements has a first array of sensor elements spatially distributed relative to a first axis and a second array of sensor elements spatially distributed relative to a second axis extending transversely to said first axis,

said reference means has means for storing a first reference signal and a second reference signal representative of frequency magnitudes and phase angles of one of said frequency-dependent signals generated by said first array and said second array respectively, and

said delay estimator means has means for generating, a first delay signal and a second delay signal representative of the time delay between said first reference signal and a frequency-dependent signal generated by said first array and said second reference signal and a frequency-dependent signal generated by said second array, and means for generating a position signal, as a function of said first delay signal and said second delay signal, representative of the position of said detected signal relative to said first and second arrays.

* * * * *

CODEN: JASMAN

The Journal of the Acoustical Society of America

ISSN: 0001-4966

Vol. 110, No. 6 December 2001

(Continued)

THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA VOL. 110, NO. 6, DECEMBER 2001

CONTENTS—*Continued from preceding page*

(Continued)

THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA VOL. 110, NO. 6, DECEMBER 2001

CONTENTS—*Continued from preceding page*

(Continued)

THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA VOL. 110, NO. 6, DECEMBER 2001

CONTENTS—*Continued from preceding page*

Document Delivery: Copies of journal articles can be ordered from *DocumentStore*, our online document delivery service (URL: http://documentstore.org/).

A two-microphone dual delay-line approach for extraction of a speech sound in the presence of multiple interferers[a)]

Chen Liu,[b)] Bruce C. Wheeler, William D. O'Brien, Jr., Charissa R. Lansing, Robert C. Bilger, Douglas L. Jones, and Albert S. Feng[c)]
Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801

(Received 28 June 2000; revised 14 June 2001; accepted 19 September 2001)

This paper describes algorithms for signal extraction for use as a front-end of telecommunication devices, speech recognition systems, as well as hearing aids that operate in noisy environments. The development was based on some independent, hypothesized theories of the computational mechanics of biological systems in which directional hearing is enabled mainly by binaural processing of interaural directional cues. Our system uses two microphones as input devices and a signal processing method based on the two input channels. The signal processing procedure comprises two major stages: (i) source localization, and (ii) cancellation of noise sources based on knowledge of the locations of all sound sources. The source localization, detailed in our previous paper [Liu *et al.*, J. Acoust. Soc. Am. **108**, 1888 (2000)], was based on a well-recognized biological architecture comprising a dual delay-line and a coincidence detection mechanism. This paper focuses on description of the noise cancellation stage. We designed a simple subtraction method which, when strategically employed over the dual delay-line structure in the broadband manner, can effectively cancel multiple interfering sound sources and consequently enhance the desired signal. We obtained an 8–10 dB enhancement for the desired speech in the situations of four talkers in the anechoic acoustic test (or 7–10 dB enhancement in the situations of six talkers in the computer simulation) when all the sounds were equally intense and temporally aligned. [DOI: 10.1121/1.1419090]

PACS numbers: 43.72.Ar, 43.72.Dv, 43.60.Bf [DOS]

I. INTRODUCTION

Selective hearing is a useful mechanism for extracting desired signals in complex acoustic environments such as a cocktail party. This so-called "cocktail party" effect has been studied psychophysically (Cherry, 1953; Blauert, 1983; Bregman, 1990; Bronkhorst and Plomp, 1992). The ability to hear in complex acoustic environments is largely attributed to the capacity to discern the spatial origins of sound sources. The neural circuitry and the underlying mechanisms for sound localization are fairly well established (Konishi *et al.*, 1988; Takahashi and Keller, 1994; Yin and Chan, 1990). Sound localization involves binaural processing of minute differences in time, intensity, and spectrum between the two ears. However, although we know the capacity of the auditory system to selectively attend to sounds originating from one source and suppress the other sounds in the ambiance, the underlying mechanisms for doing so are largely unknown. Therefore, designing an artificially intelligent system today to achieve selective hearing is still largely based on our relatively rich knowledge of the physical world (e.g., signal processing techniques) plus our limited knowledge of the biological world.

One of the prominent noise suppression concepts is the EC or equalization-and-cancellation scheme of Durlach (Durlach, 1960, 1972). It requires two inputs followed by a two-stage signal processing: (i) equalization that makes the noise components identical in both channels; and (ii) cancellation or subtraction of the noise components in one channel from those in the other channel. Actually most two-microphone-based noise cancellation techniques to date (e.g., Widrow *et al.*, 1975; Strube, 1981; Chabries *et al.*, 1982; Chazan *et al.*, 1988; Weiss, 1987; Peterson *et al.*, 1987) are essentially variants of the EC scheme and differ primarily in the procedures by which the filter parameters are adapted. Thus far, these have rendered satisfactory noise reduction only for situations in which there are one desired source and one noise source.

Our noise cancellation technique described herein also falls in this category. However, it is devised so as to cancel multiple noise sources more efficiently by capitalizing on the knowledge of the spatial directions of the sound sources in the environment. For the purpose of sound localization, we have designed a system (Liu *et al.*, 2000) based on a broadband "dual delay-line" structure and the coincidence detection principle of Jeffress (1948). Our noise cancellation technique also adopts the dual delay-line as the infrastructure.

So far the Jeffress model has been studied and various modifications have been developed to account for different psychological observations (see reviews in the book chapters by Colburn and Durlach, 1978; Colburn, 1996; and Stern and Trahiotis, 1995, 1997). It was only recently that the Jeffress model began to be considered for use in the extraction of

---

[a)]Portions of this paper were presented at the Hearing Aid Research & Development Conference, Bethesda, MD, September 1997.

[b)]Present address: Motorola Labs, 1141 Opus Place, Downers Grove, IL 60515.

[c)]Electronic mail: a-feng@uiuc.edu

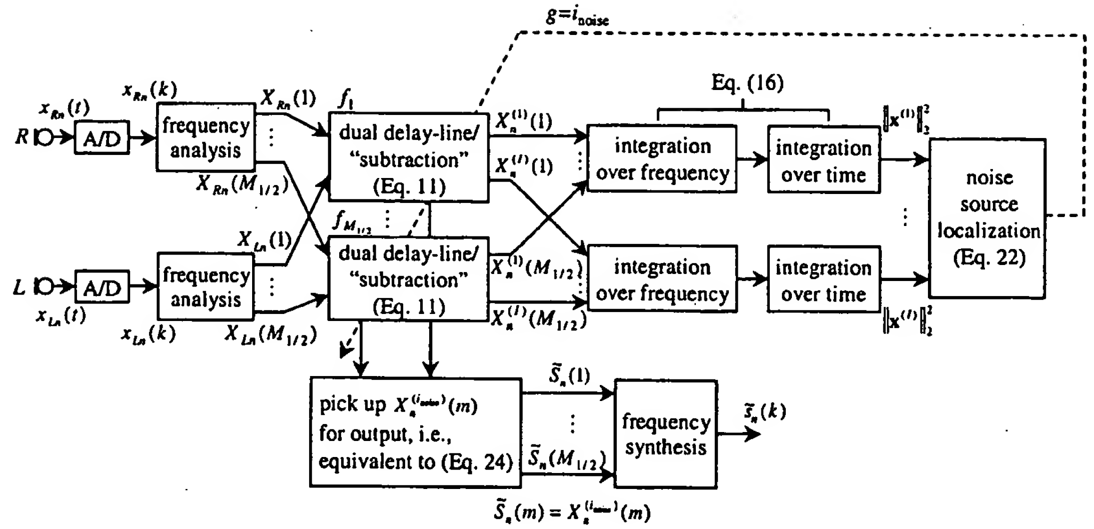


FIG. 1. The block diagram of System I for extraction of the desired source, whose location is known *a priori*, in the presence of one noise source whose location is estimated by the system.

speech in noise (e.g., Bodden, 1993; Banks, 1993). Since the model maps the acoustic space into a network, one appealing feature is the potential for detecting the number and the azimuths of sound sources present in the auditory space. This provides the mechanism by which the system can focus on, and extract the signal from, one desired source direction, while at the same time suppressing the sounds arising from the other directions. In his acoustic processor, Bodden (1993, 1996) basically took the Jeffress' coincidence sound localization models, as implemented by Lindemann (1986) and Gaik (1993), and added a time-variant Wiener filter for noise cancellation after sounds had been localized. However, since it is impossible to obtain an accurate estimate of the power density spectra of both the desired and noise signals, the result will always have residual noise and some cancellation and distortion of the desired signal.

The work described herein was motivated by the need to find a general solution for signal extraction in real world situations where there are multiple (>2) concurrent sound sources. Our signal extraction technique evolved from a subtraction procedure. Note that, interestingly, subtraction is also employed in the directional hearing mechanism with a pressure-gradient receiver (Feng and Shofner, 1981). Theoretically, a conventional noise cancellation system using a two-microphone array performs well when there are two sources but its performance degrades rapidly as the number of sources increases. To attack this problem, we developed a broadband noise cancellation strategy, making the two-microphone array subtraction approach more effective by taking advantage of the dual delay-line structure.

In this paper, we first introduce a subtraction method, which is the core of our noise cancellation technique. The subtraction procedure is then extended via the broadband dual delay-line structure for cancellation of multiple sources. In Sec. II A, we describe the subtraction procedure in the context of extracting a desired source at a known location in the presence of one interfering source at an arbitrary location. The subtraction operation is mathematically analyzed in Sec. II B. Section II C gives a beamforming interpretation of the subtraction method. In Sec. II D, the method is generalized to situations in which neither the location of the desired source nor that of the interference is known. Section III describes a strategy for extending the method to a system suitable for cancellation of multiple interfering sources. Section IV presents the experimental results and analysis. Discussion of several practical issues is given in Sec. V.

II. INTRODUCTION TO THE NEW CANCELLATION SCHEME

A. Cancellation algorithm based on the dual delay-line structure

In this section we will describe a new noise cancellation algorithm. It is fundamentally a subtraction operation applied on the two input signals. The signals are received by two microphones, which are paired with a fixed inter-microphone distance. The subtraction is conducted based on the infrastructure of the dual delay-line network in the frequency domain. A block diagram of the basic signal processing system (System I) is shown in Fig. 1. The two inputs, $x_{Ln}(t)$ and $x_{Rn}(t)$, are digitized, their digital versions being $x_{Ln}(k)$ and $x_{Rn}(k)$, respectively. Their spectra, $X_{Ln}(m)$ and $X_{Rn}(m)$, $m=1,\ldots,M$, are obtained through discrete Fourier transform (DFT). The subscripts L and R denote left and right channels, and n the frame index of the short-term Fourier analysis.

For clarity, we shall focus on the system description for an arbitrary frequency ω_m. For each frequency, the complex signals from the two channels are fed into a pair of delay-lines (Fig. 2), both of which are composed of I delay units with delay values τ_i $(i=1,\ldots,I)$ given by

$$\tau_i=\frac{\mathrm{ITD}_{\max}}{2}\sin\left(\frac{i-1}{I-1}\pi-\frac{\pi}{2}\right),\quad i=1,\ldots,I, \tag{1}$$

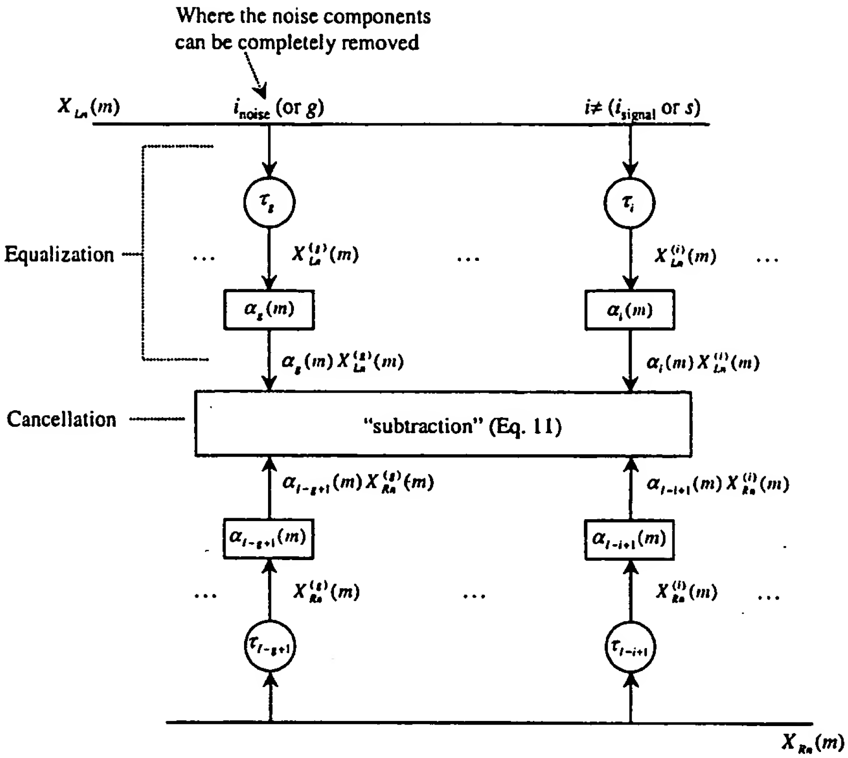


FIG. 2. The dual delay-line used as the basic structure in our system.

where $\mathrm{ITD}_{\max}=D/c$ is the maximum inter-microphone time difference, D is the inter-microphone distance, c is the speed of sound, and I is an odd number greater than 1. By using Eq. (1), it can be shown that the time delays are antisymmetric with respect to the midpoint $(I+1)/2$, i.e.,

$$\tau_{I-i+1}=-\tau_i, \quad i=1,\ldots,(I+2)/2-1. \tag{2}$$

It is noted that if there is no diffracting object between the two microphones, the horizontal plane can be uniformly divided into I sectors with the azimuth of each sector being

$$\theta_i=\frac{\pi}{2}-\frac{i-1}{I-1}\pi, \quad i=1,\ldots,I. \tag{3}$$

Therefore, the azimuths can be mapped one-to-one onto the corresponding positions in the delay line

$$\tau_i=-\frac{\mathrm{ITD}_{\max}}{2}\sin\theta_i, \quad i=1,\ldots,I. \tag{4}$$

Note that the resolution of the dual delay-lines representing the spatial azimuth is determined by the values of the time-delay units τ_i that have a time unit such as millisecond. As will be described in detail next, the delays are applied to the left and right input signals at each frequency in the frequency domain. Thus the dual delay-line works like rotating two separate phasors in the opposite directions in the complex plane [Eqs. (9) and (10)] until they are in phase, i.e., the so-called coincidence operation. The step size of the rotation, $\exp(-j\omega_m\tau_i)$, can be arbitrarily small in the frequency domain. Therefore, the azimuthal resolution of the dual delay-lines is not controlled by the sampling rate. Some other relevant discussions will be given in Sec. II C.

Figure 2 shows that the dual delay-line structure is similar to that adopted previously for sound localization (Liu *et al.*, 2000) except that a compensation element $\alpha_i(m)$ has been added following each delay unit. These elements, which compensate for differences in the intensity of noise at the two microphones, are functions of both azimuth and frequency. Appendix A derives the compensation values for the ideal case of point sources with distance-dependent amplitude decline, in a lossless medium. In practice, however, all the values of $\alpha_i(m)$ and τ_i $(i=1,\ldots,I)$ are to be adjusted empirically the same time when the system is being calibrated to compensate for asymmetries between the two microphones. The compensation factors remain fixed so long as the asymmetries are not changed. This fixed interaural intensity difference (IID) corresponding to each interaural time difference (ITD) mimics that observed in humans (Gaik, 1993). In the anechoic chamber tests reported below, the values of ITD units τ_i $(i=1,\ldots,I)$ were set uniformly while the values of IID $\alpha_i(m)$ $(i=1,\ldots,I)$ were determined empirically.

In this subsection let us suppose the direction of the desired source is known *a priori* and we use $i_{\mathrm{signal}}=s$ to denote the in-phase (coincident) position along the dual delay-line for the desired signal components. We use $i_{\mathrm{noise}}=g$ to denote the in-phase position for the noise signal components. Note that the position index along the dual delay-line is coincident with the index of the delay units in the left channel. After equalization, the in-phase desired signal components are identical in the left and right channels at $i_{\mathrm{signal}}=s$, which is assumed to be $S_n(m)=A_s\exp[j(\omega_m t+\phi_s)]$; likewise, the in-phase noise signal component is identical in the left and right channels at $i_{\mathrm{noise}}=g$, which is assumed to be $G_n(m)=A_g\exp[j(\omega_m t+\phi_g)]$, where ϕ_s and ϕ_g are the initial phases for signal and noise, respectively. Based on

these assumptions, the left and right channel input (microphone) signals are, respectively,

$$X_{Ln}(m)=\frac{1}{\alpha_s(m)}S_n(m)\exp(j\omega_m\tau_s)+\frac{1}{\alpha_g(m)}G_n(m)\exp(j\omega_m\tau_g) \quad (5)$$

and

$$X_{Rn}(m)=\frac{1}{\alpha_{I-s+1}(m)}S_n(m)\exp(j\omega_m\tau_{I-s+1})+\frac{1}{\alpha_{I-g+1}(m)}G_n(m)\exp(j\omega_m\tau_{I-g+1}). \quad (6)$$

Then, we can find the mathematical representation for the equalized signals $\alpha_i(m)X_{Ln}^{(i)}(m)$ for the left channel, and $\alpha_{I-i+1}(m)X_{Rn}^{(i)}(m)$ for the right channel at any arbitrary point i (except $i=s$), along the dual delay-line. They are

$$\alpha_i(m)X_{Ln}^{(i)}(m)=\frac{\alpha_i(m)}{\alpha_s(m)}S_n(m)\exp[j\omega_m(\tau_s-\tau_i)]+\frac{\alpha_i(m)}{\alpha_g(m)}G_n(m)\exp[j\omega_m(\tau_g-\tau_i)] \quad (7)$$

and

$$\alpha_{I-i+1}(m)X_{Rn}^{(i)}(m)=\frac{\alpha_{I-i+1}(m)}{\alpha_{I-s+1}(m)}S_n(m)\times\exp[j\omega_m(\tau_{I-s+1}-\tau_{I-i+1})]+\frac{\alpha_{I-i+1}(m)}{\alpha_{I-g+1}(m)}G_n(m)\times\exp[j\omega_m(\tau_{I-g+1}-\tau_{I-i+1})], \quad (8)$$

where

$$X_{Ln}^{(i)}(m)=X_{Ln}(m)\exp(-j\omega_m\tau_i) \quad (9)$$

and

$$X_{Rn}^{(i)}(m)=X_{Rn}(m)\exp(-j\omega_m\tau_{I-i+1}). \quad (10)$$

The subtraction step in the algorithm performs the following operation on each signal pair, $\alpha_i(m)X_{Ln}^{(i)}(m)$ and $\alpha_{I-i+1}(m)X_{Rn}^{(i)}(m)$, for $i=1,\ldots,I$, at any location along the delay line except the location where $i=s$:

$$X_n^{(i)}(m)=\frac{\alpha_i(m)X_{Ln}^{(i)}(m)-\alpha_{I-i+1}(m)X_{Rn}^{(i)}(m)}{[\alpha_i(m)/\alpha_s(m)]\exp[j\omega_m(\tau_s-\tau_i)]-[\alpha_{I-i+1}(m)/\alpha_{I-s+1}(m)]\exp[j\omega_m(\tau_{I-s+1}-\tau_{I-i+1})]}, \quad \text{for } i\neq s. \quad (11)$$

A caveat in using Eq. (11) is that if the value of the denominator is too small, a small positive constant ϵ is added to limit the magnitude of $X_n^{(i)}(m)$.

B. Physical meaning of the delay-line subtraction operation

To analyze the operation, Eq. (11) can be expressed in the following form via substitution of Eqs. (7) and (8):

$$X_n^{(i)}(m)=S_n(m)+G_n(m)v_{s,g}^{(i)}(m), \quad i\neq s, \quad (12)$$

where

$$v_{s,g}^{(i)}(m)=\frac{[\alpha_i(m)/\alpha_g(m)]\exp[j\omega_m(\tau_g-\tau_i)]-[\alpha_{I-i+1}(m)/\alpha_{I-g+1}(m)]\exp[j\omega_m(\tau_{I-g+1}-\tau_{I-i+1})]}{[\alpha_i(m)/\alpha_s(m)]\exp[j\omega_m(\tau_s-\tau_i)]-[\alpha_{I-i+1}(m)/\alpha_{I-s+1}(m)]\exp[j\omega_m(\tau_{I-s+1}-\tau_{I-i+1})]}, \quad i\neq s. \quad (13)$$

Equations (11) and (13) can be simplified when the antisymmetric relationship in Eq. (2) is used. Thus,

$$X_n^{(i)}(m)=\frac{\alpha_i(m)X_{Ln}^{(i)}(m)-\alpha_{I-i+1}(m)X_{Rn}^{(i)}(m)}{[\alpha_i(m)/\alpha_s(m)]\exp[j\omega_m(\tau_s-\tau_i)]-[\alpha_{I-i+1}(m)/\alpha_{I-s+1}(m)]\exp[-j\omega_m(\tau_s-\tau_i)]}, \quad \text{for } i\neq s, \quad (14)$$

and

$$v_{s,g}^{(i)}(m)=\frac{[\alpha_i(m)/\alpha_g(m)]\exp[j\omega_m(\tau_g-\tau_i)]-[\alpha_{I-i+1}(m)/\alpha_{I-g+1}(m)]\exp[-j\omega_m(\tau_g-\tau_i)]}{[\alpha_i(m)/\alpha_s(m)]\exp[j\omega_m(\tau_s-\tau_i)]-[\alpha_{I-i+1}(m)/\alpha_{I-s+1}(m)]\exp[-j\omega_m(\tau_s-\tau_i)]}, \quad i\neq s. \quad (15)$$

When ignoring the compensation factors $\alpha_i(m)$, an interesting observation of the subtraction [Eq. (11) or (14)] is that it computes the difference between each pair of taps at the ith location divided (shifted) by a factor that is determined only by the difference in time delay between that location and the location corresponding to the desired signal. Next we will show that Eq. (11) performed at the location i in the dual delay-line corresponding to the noise source will cancel the noise signal and provide an estimate of the desired signal. Moreover, the location can be found using an energy quantity.

A signal vector containing all the frequency components for the preceding N time frames is $\mathbf{x}^{(i)}=(X_1^{(i)}(1), X_1^{(i)}(2),\ldots,X_1^{(i)}(M),X_2^{(i)}(1),\ldots,X_2^{(i)}(M),\ldots,X_N^{(i)}(1),\ldots,X_N^{(i)}(M))^T$, $i=1,\ldots,I$, where T denotes vector transposition. The energy $E[\mathbf{x}^{(i)}]$ of vector $\mathbf{x}^{(i)}$ is

$$\begin{aligned} E[\mathbf{x}^{(i)}]=\|\mathbf{x}^{(i)}\|_2^2&=\sum_{n=1}^{N}\sum_{m=1}^{M}|X_n^{(i)}(m)|^2\\ &=\sum_{n=1}^{N}\sum_{m=1}^{M}|S_n(m)+G_n(m)v_{s,g}^{(i)}(m)|^2,\\ &\quad i=1,\ldots,I, \end{aligned} \tag{16}$$

where the energy of the signal $X_n^{(i)}(m)$ is

$$E[X_n^{(i)}(m)]=|X_n^{(i)}(m)|^2=|S_n(m)+G_n(m)v_{s,g}^{(i)}(m)|^2. \tag{17}$$

To separate the complex signal into the desired signal and noise, we define the following vectors in the similar manner

$$\mathbf{s}=(S_1(1),S_1(2),\ldots,S_1(M),S_2(1),\ldots,S_2(M),\ldots,S_N(1),\ldots,S_N(M))^T,$$

and

$$\begin{aligned}\mathbf{g}^{(i)}=(&G_1(1)v_{s,g}^{(i)}(1),G_1(2)v_{s,g}^{(i)}(2),\ldots,\\ &G_1(M)v_{s,g}^{(i)}(M),G_2(1)v_{s,g}^{(i)}(1),\ldots,\\ &G_2(M)v_{s,g}^{(i)}(M),\ldots,G_N(1)v_{s,g}^{(i)}(1),\ldots,\\ &G_N(M)v_{s,g}^{(i)}(M))^T,\end{aligned}$$

where $i=1,\ldots,I$. The energy of $\mathbf{s}$ and $\mathbf{g}^{(i)}$ are, respectively,

$$E[\mathbf{s}]=\|\mathbf{s}\|_2^2=\sum_{n=1}^{N}\sum_{m=1}^{M}|S_n(m)|^2 \tag{18}$$

and

$$E[\mathbf{g}^{(i)}]=\|\mathbf{g}^{(i)}\|_2^2=\sum_{n=1}^{N}\sum_{m=1}^{M}|G_n(m)v_{s,g}^{(i)}(m)|^2,\quad i=1,\ldots,I. \tag{19}$$

In general, the desired signal and the noise signal are independent. Thus, vectors $\mathbf{s}$ and $\mathbf{g}^{(i)}$ are orthogonal. According to the Pythogoras Theorem, we would have the following relationship:

$$\begin{aligned} E[\mathbf{x}^{(i)}]&=\|\mathbf{x}^{(i)}\|_2^2\\ &=\|\mathbf{s}+\mathbf{g}^{(i)}\|_2^2\\ &=\|\mathbf{s}\|_2^2+\|\mathbf{g}^{(i)}\|_2^2=E[\mathbf{s}]+E[\mathbf{g}^{(i)}],\quad i=1,\ldots,I. \end{aligned} \tag{20}$$

Because $\|\mathbf{g}^{(i)}\|_2^2\geqslant 0$,

$$E[\mathbf{x}^{(i)}]=\|\mathbf{x}^{(i)}\|_2^2\geqslant\|\mathbf{s}\|_2^2=E[\mathbf{s}],\quad i=1,\ldots,I. \tag{21}$$

The equality in Eq. (21) is satisfied, or equivalently $\min E[\mathbf{x}^{(i)}]$ occurs, only when $E[\mathbf{g}^{(i)}]=\|\mathbf{g}^{(i)}\|_2^2=0$, which happens in either of the following two conditions:

(a) When $G_n(m)=0$, i.e., the noise source is silent. In this case, there is no need for doing localization of the noise source and noise cancellation.

(b) When $v_{s,g}^{(i)}(m)=0$, Eq. (15) indicates that this case corresponds to $i=g=i_{\text{noise}}$. Therefore, $E[\mathbf{x}^{(i)}]$ has its minimum at $i=g=i_{\text{noise}}$ and the minimum value, according to Eq. (21), is $E[\mathbf{s}]$. Thus,

$$E[\mathbf{s}]=E[\mathbf{x}^{(i_{\text{noise}})}]=\min_i E[\mathbf{x}^{(i)}]. \tag{22}$$

When $i=i_{\text{noise}}$, Eq. (12) provides

$$\begin{aligned}\tilde{S}_n(m)&=X_n^{(i_{\text{noise}})}(m)\\ &=S_n(m)+G_n(m)v_{s,g}^{(i_{\text{noise}})}(m)=S_n(m).\end{aligned} \tag{23}$$

In other words, in the presence of one desired source and one noise source, the subtraction operation [Eq. (11)] applied at the location $i=g(=i_{\text{noise}})$ in the dual delay-line structure can produce an accurate estimate of the desired signal. Namely,

$$X_n^{(g)}(m)=\frac{\alpha_g(m)X_{Ln}^{(g)}(m)-\alpha_{I-g+1}(m)X_{Rn}^{(g)}(m)}{[\alpha_g(m)/\alpha_s(m)]\exp[j\omega_m(\tau_s-\tau_g)]-[\alpha_{I-g+1}(m)/\alpha_{I-s+1}(m)]\exp[j\omega_m(\tau_{I-s+1}-\tau_{I-g+1})]}. \tag{24}$$

The above analysis with energy also suggests a simple method to estimate the location $g=i_{\text{noise}}$ of the noise source in the two-source situation where the direction of the desired source is known *a priori*. Specifically, localization of the noise source can be conducted by finding the location i_{noise} along the dual delay-line that produces the minimum value of $E[\mathbf{x}^{(i)}]$ [Eqs. (11), (16), and (22)]. Once the location i_{noise} is determined, the azimuth of the noise source is easily determined by using Eq. (3). The estimated noise location i_{noise} can be fed back to the dual delay-line for noise cancellation

and extraction of the desired signal using Eq. (24).

In Fig. 1, the blocks labeled "Integration over Time" and "Integration over Frequency" together calculate the energy $E[\mathbf{x}^{(i)}]$ defined in Eq. (16). The block labeled "Noise Source Localization" locates the minimum point of $E[\mathbf{x}^{(i)}]$, and then supplies this as the estimate of i_{noise} to the dual delay-line. Since all the components $X_n^{(i)}(m)$, $i=1,\ldots,I$, have been computed at the localization step, now we only need to pick up the appropriate component $X_n^{(i_{\text{noise}})}(m)$, i.e., $\tilde{S}_n(m)$; Eq. (24) does not need to be actually executed in this case. Note that all the frequency computations so far are conducted on the first half $(m=1,\ldots,M_{1/2})$ of the whole bandwidth. The block labeled "Frequency Synthesis" derives the second half $(m=M_{1/2}+1,\ldots,M)$ by means of the symmetry property of the inverse discrete Fourier transform (IDFT) and then conducts the IDFT to generate the time-domain signal $\tilde{s}_n(k)$.

C. Beamforming interpretation of the delay-line subtraction operation

This system can be understood conceptually by an equivalent beamforming procedure. Equation (11) can be expressed in the following form:

$$X_n^{(i)}(m)=w_{Ln}^{(i)}(m)X_{Ln}^{(i)}(m)+w_{Rn}^{(i)}(m)X_{Rn}^{(i)}(m), \tag{25}$$

where $w_{Ln}^{(i)}(m)$ and $w_{Rn}^{(i)}(m)$ are beamforming weights. That is, for each location along the dual delay-line at each frequency, a specific nulling pattern is generated with the null pointed toward the direction corresponding to the delay-line location while the gain in the presumed target direction is kept unity. Figure 3(a) shows a broadband intelligibility-weighted beampattern (for definition, see Appendix B) for selected nulling directions at −80°, −60°, −40°, −20°, 20°, 40°, 60°, and 80° azimuth (labeled A through H, respectively) with the desired source at 0° azimuth. It can be seen that Eq. (11) actually positions a null in the direction of the noise source while keeping the broadband gain always unity in the direction of the desired source. Since each nulling pattern uses only 2 degrees of freedom, i.e., $w_{Ln}^{(i)}(m)$ and $w_{Rn}^{(i)}(m)$, to satisfy the two constraints (directions of null and unity-gain), the null patterns are fixed and there is no room to play optimization on the pattern shape. As will be presented in Sec. III, this study, by taking advantages of the dual delay-line network, the estimated source locations, as well as the broadband characteristics of dialog speech, sought to find an appropriate strategy [which is a nonlinear one as shown in Eq. (27)] to utilize the simple null patterns for target extraction among multiple interferers.

To extend the discussion on the azimuthal resolution of the dual delay-lines in Sec. II A, let us look at a distinguished feature of Eq. (11). In the numerator of Eq. (11), the signals in the two channels can be phase-shifted by any arbitrary (small) values in the frequency domain. However, the denominator eliminates the effect and thus $X_n^{(i)}(m)$ contains an intact component of the desired signal $S_n(m)$. Moreover, at the location $i=g=i_{\text{noise}}$ where the noise component is completely cancelled, only the desired signal is left in the result. If interpreted as a beamformer, Eq.

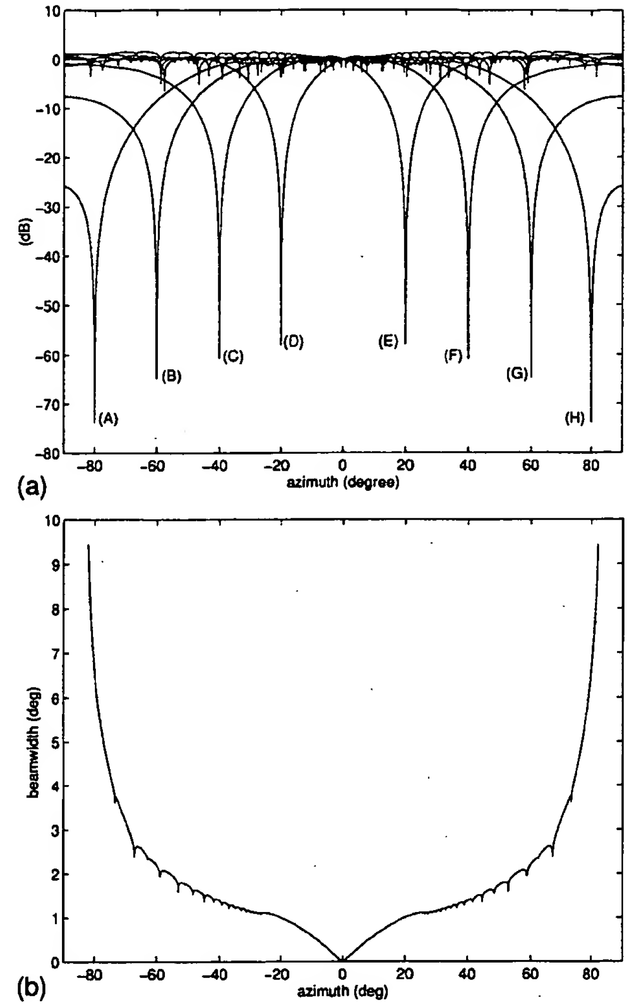


FIG. 3. (a) The intelligibility-weighted beampattern created by Eq. (11) for the cases where the desired source was always at 0° azimuth while the noise source was at −80° (A), −60° (B), −40° (C), −20° (D), 20° (E), 40° (F), 60° (G), and 80° (H) azimuth, respectively. The inter-microphone distance in this example was 144 mm. (b) The null-width of the intelligibility-weighted beampattern at −30 dB as a function of azimuth.

(11) operated on the dual delay-line in the frequency domain enables a null steering precisely in any arbitrary direction regardless of the sampling rate.

D. Extended application

The method suggested in the preceding subsection for localization and cancellation of the noise source is valid only when the direction of the desired source is known *a priori*. It cannot be directly applied in the situations where the direction of the desired source is also unknown. Therefore, we designed another system (System II in Fig. 4). The operation of this system is divided into two steps: it localizes both the desired source and noise source, and then selectively extracts the signal from the desired direction. The localization step employs an efficient localization method comprising dual delay-line coincidence detection followed by a nonlinear operation and then temporal and spectral integrations. The method was described in detail in a previous paper (Liu *et al.*, 2000) in which it was shown to accurately localize

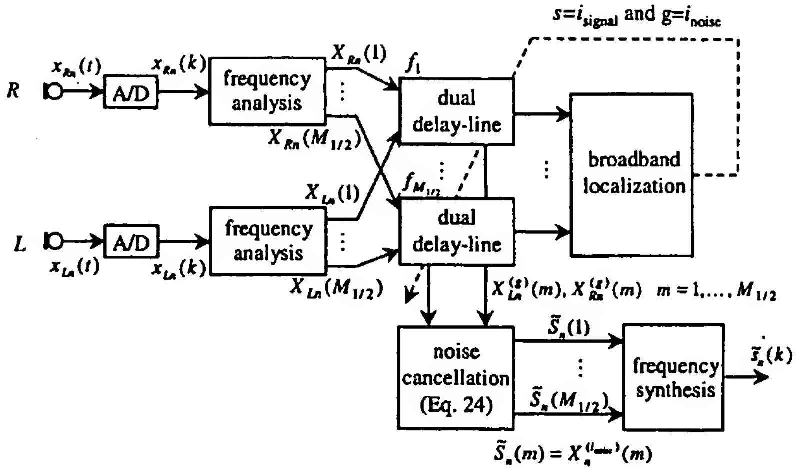


FIG. 4. The block diagram of System II for extraction of one desired source in the presence of a noise source when both source locations have to be estimated by the system. See Liu *et al.* (2000) for details about the block "Broadband Localization."

four talkers in an anechoic room and six talkers in computer simulation. Thus the localization block in Fig. 4 determines the in-phase positions, $i_{\mathrm{signal}}=s$ and $i_{\mathrm{noise}}=g$, of both the desired and noise sources along the dual delay-line, which were then used by the subtraction in Eq. (24) for extracting the desired signal $\tilde{S}_n(m)$. That is, except for the separate source localization step, System II employs the same noise cancellation method as described in the preceding subsection [Eq. (24)].

In comparison with System I, System II (without the assumption of direction of the desired source) is functionally more flexible. For example, in a situation with two talkers, there is no need to align the dual microphones physically toward one talker, and either talker can be taken as the desired one. The user can choose between the two sources at any time by using an electronic switch instead of changing the pointing direction of the microphones. Actually the microphones do not necessarily point to either of the sources.

We presented System I in the preceding subsection mainly for illustrating the mechanism of the dual delay-line subtraction [Eq. (11)] and shows its capacity for both noise-localization and desired-extraction. However, the operation in System I is computationally expensive because Eq. (11) must be applied to each tap in the dual delay-line for localization. Moreover, its use is limited to a two-talker (with the direction of the desired talker known *a priori*) situation. In comparison, the coincidence detection scheme for localization employed by System II is simpler in computation. What is more important is that, as we will show in the next section, System II can be further extended to situations with multiple interfering talkers.

Although our localization method worked quite well in a multiple-source environment, we normally observed relatively larger and more frequent localization errors for the lateral sources (Liu *et al.*, 2000). The robustness of the noise cancellation to localization errors can be roughly estimated by looking at the null-width of the nulls in Fig. 3(a). For example, the null-width evaluated at -30 dB is shown in Fig. 3(b). It shows that the width is wider when the direction of the null is farther away from the midline; that is, the noise cancellation method can tolerate bigger localization errors for lateral sources. Therefore, the greater localization errors for lateral sources do not degrade the system performance in terms of noise cancellation.

III. STRATEGY FOR BROADBAND MULTIPLE-SOURCE CANCELLATION

The greatest challenge associated with extension of the noise cancellation method from two-source situations to multiple-source (>2) situations is that a two-input system in theory can only effectively cancel the sound from *one* interfering source. This is due to the fact that only one null can be generated in the beampattern when using a two-microphone array. In the narrow-band situation, an apparent solution is to adaptively steer the null toward the most intense noise source at each moment. In the broadband situation, since the input signal is decomposed into its frequency components, one can formulate a separate one-null beampattern for each frequency. When there is one noise source as described in the preceding section, the nulls at all frequencies are steered in the same direction of the single noise source. However, when there are more than two sources, each frequency bin can be treated separately so that its beampattern null is adaptively steered at each moment toward the noise source that is emitting the most intense energy at that frequency, while maintaining unity gain toward the desired source. It is actually a dynamic application of the subtraction operation in Eq. (24). This noise cancellation strategy is based on the following rationale:

(a) Natural speech has many pauses and silent intervals, both of which usually occupy 60%–65% of the total time (Flanagan, 1972, p. 386). Therefore, when multiple talkers speak simultaneously, there are always a number of short temporal gaps present. The number of overlapping talkers at each moment is usually smaller than the total number of talkers.

(b) Even when multiple talkers speak at the same moment, different talkers likely dominate at different frequency bins at each moment due to the differences in articulation such as voicing and pitch. There are about ten phonemes per second in conversational speech, more than 60% of which are low-energy, high-frequency consonants, and less than 40% of which are high-

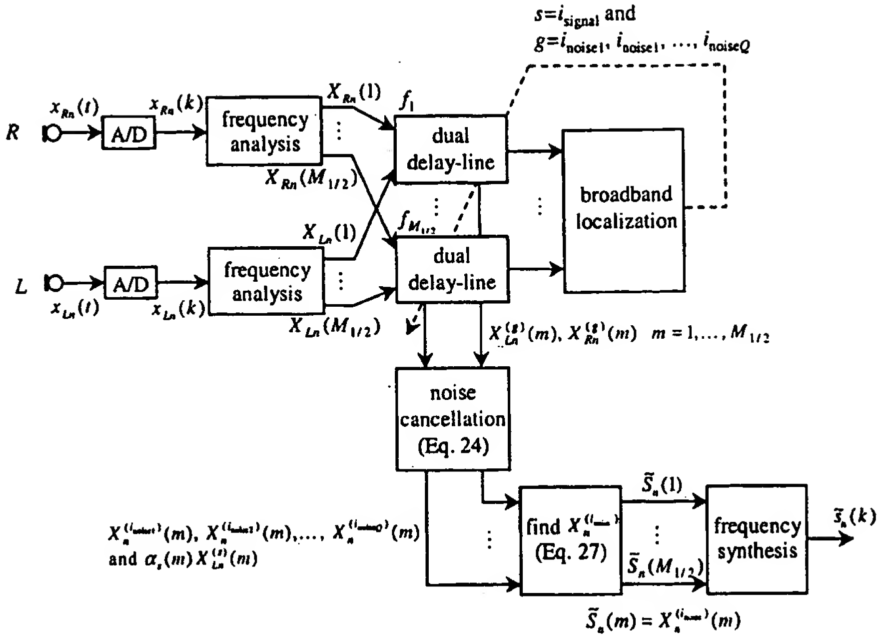


FIG. 5. The block diagram of System III for extraction of one desired source in the presence of more than one noise source when all the source locations have to be estimated by the system. See Liu *et al.* (2000) for details about the block "Broadband Localization."

energy, low-frequency vowels (Flanagan, 1972, p. 4). In the presence of multiple talkers, the talkers who are articulating the high-energy vowels are dominant in both the localization (Liu *et al.*, 2000) and cancellation [Eq. (27)] and hence more easily removed. On the other hand, due to the asymmetry of the filtering response of the human ear, the masking effect of low frequencies on high frequencies is much stronger than the reverse (Jeffress, 1970, p. 95). In other words, cancellation of a talker at his/her strongest frequency components, which are likely the major components of a vowel, may effectively cancel the masking effect of the talker.

To obtain the information about location of each source for the noise cancellation algorithm, the localization algorithm in Liu *et al.* (2000) is employed. Suppose there are Q noise sources with corresponding locations in the dual delay-line being $i_{\mathrm{noise1}}, i_{\mathrm{noise2}}, \ldots, i_{\mathrm{noise}Q}$. By applying Eq. (24), we obtain $X_n^{(i_{\mathrm{noise1}})}(m), X_n^{(i_{\mathrm{noise2}})}(m), \ldots, X_n^{(i_{\mathrm{noise}Q})}(m)$ for each frequency ω_m. If the localization is accurate, they all should include the component of the desired signal at frequency ω_m as well as components from interfering sources other than the one to be canceled. In order to determine the particular noise source to be canceled, the energies of $X_n^{(i_{\mathrm{noise1}})}(m), X_n^{(i_{\mathrm{noise2}})}(m), \ldots, X_n^{(i_{\mathrm{noise}Q})}(m)$ are calculated and compared. The minimum $X_n^{(i_{\mathrm{noise}})}(m)$ is taken as output $\tilde{S}_n(m)$:

$$\tilde{S}_n(m)=X_n^{(i_{\mathrm{noise}})}(m), \tag{26}$$

where $X_n^{(i_{\mathrm{noise}})}(m)$ satisfies the following condition:

$$|X_n^{(i_{\mathrm{noise}})}(m)|^2=\min\{|X_n^{(i_{\mathrm{noise1}})}(m)|^2,\; |X_n^{(i_{\mathrm{noise2}})}(m)|^2,\ldots,|X_n^{(i_{\mathrm{noise}Q})}(m)|^2, |\alpha_s(m)X_{Ln}^{(s)}(m)|^2\}. \tag{27}$$

By referring to the energy analysis in Sec. II B, it is easy to see that this strategy is logically consistent with Eq. (22). It is noted that, in Eq. (27), we also include the original signal $\alpha_s(m)X_{Ln}^{(s)}(m)$ for the following reason. The beampattern designed above sometimes may amplify other less intense noise sources. When the amount of noise amplification is larger than the amount of cancellation of the most intense noise source, it may be better to keep the input signal at that frequency at that moment unchanged. An extended system (System III in Fig. 5) was developed using System II (Fig. 4) as the foundation. In comparison with System II, it identifies multiple (>2) source directions and tentatively cancels each noise source; specifically it cancels the instantaneously most intense source on a frequency-by-frequency basis [Eq. (27)].

The cancellation step relies on the localization step to provide azimuth information for each source, which is usually a difficult task especially in the presence of multiple sources. However, as shown in our previous paper (Liu *et al.*, 2000), our localization system can satisfactorily localize four sources in an anechoic room and six sources in simulation, if not more. In addition, the cancellation step does not have rigid requirement that all the sources must be localized accurately. As a matter of fact, our strategy is to cancel the strongest noise component at each frequency bin—this is usually emitted from one of those momentarily relatively intense noise sources, which are easy to localize compared with other relatively less intense sources.

IV. EXPERIMENT

For the case of two talkers, once the locations of the talkers are determined, the sound from one talker can be removed by using System I or System II with essentially no residual noise while the estimated desired signal is distortionless. This was clearly supported in theory and also pre-

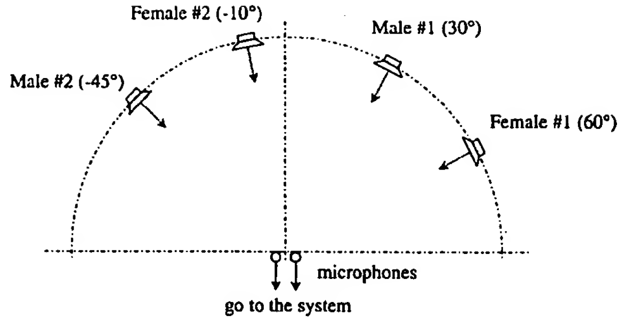


FIG. 6. Top view of the spatial configuration of one of our experimental setups. This experimental setup corresponds to test configuration #2 in Tables I and II.

viously demonstrated by Banks (1993). In this paper we present the results of four-talker experiments using the noise cancellation of System III.

The experiments employed a record-and-play procedure with four high-fidelity loudspeakers (ADS 200LC) and were conducted in an anechoic room and in a moderately quiet conference room with a reverberation time constant of approximately 400 ms. The speech materials consisted of spondaic words spoken by native speakers of American English; these were recorded in a sound studio at the Beckman Institute. All the speech recordings were equalized in average intensity and played through the loudspeakers. The words in each experimental condition were temporarily aligned, i.e., it was equivalent to all the talkers starting at the same time. The inter-microphone distance was 144 mm. All the loudspeakers were at a fixed equal distance of 1.0 m (unless otherwise stated) from the midpoint between the two microphones, and all the loudspeakers and microphones were at the same elevation ($\sim$1 m from the floor). Correspondingly, the compensation factors in Eq. (11) were determined for that distance.

The signals were low-pass filtered at 6 kHz and sampled at a 12.8-kHz rate with 16-bit quantization. In the short-term spectral analysis, a 20-ms segment of signal was weighted by a Hamming window, padded with zeros to 2048 points, and Fourier transformed with frequency resolution of about 6 Hz. Consecutive frames overlapped by 15 ms. The values of the time delay units τ_i $(i=1,\ldots,I)$ were determined such that the dual delay-line has a uniform azimuthal resolution of 0.5° (i.e., $I=361$).

Two groups of talkers were used in our tests. Each group consisted of four different talkers speaking different spondaic words. Five tests were conducted for each group; each test adopted a different azimuthal arrangement of the sources, with the separation between adjacent sources ranging 10°–75°. Figure 6 illustrates one of the configurations. Each test consisted of four subtests in which each talker was taken in turn as the desired source with all the other talkers as the noise sources. The localization of the talkers was conducted using both the "direct" and "stencil" methods in Liu *et al.* (2000).

The system performance was evaluated using an objective intelligibility-weighted measure, whose concept was first proposed by Peterson (1989) and described in detail in Liu and Sideman (1996). Specifically, we used intelligibility-weighted signal cancellation, intelligibility-weighted noise cancellation, and net intelligibility-weighted gain (see Appendix B for definition).

As mentioned above, an array of tests was conducted with a number of variables such as different talkers, different spondaic words, different azimuthal arrangements, different localization methods, and different combinations of the variables. However, it is not necessary to present all our results since most of the variables, as they turned out to be, have no statistically significant effect on the final noise cancellation performance. Specifically, the experimental results showed no statistical difference due to talkers and words spoken. It also showed no significant effect from using the "direct" method versus the "stencil" method for source localization (Liu *et al.*, 2000). Therefore, we only present the results from Group #1 with the location information derived using the "stencil" method. As mentioned above, it contained five tests corresponding to five different spatial configurations. For each test, we present result from only one (arbitrarily chosen) of the four subtests since the location of the desired source has no obvious effect on noise cancellation. Table I shows typical results chosen from the anechoic chamber test

TABLE I. Experiment results attained from the anechoic room using System III. The results shown were derived from five tests (configurations) from Group #1 including two male speakers (M1 and M2) and two female speakers (F1 and F2). The spondaic word spoken by each talker is given in italic. The values in parentheses are cancellation of the desired sources in dB. Configuration test #2 is shown in Fig. 6.

	Intelligibility-weighted signal cancellation (dB)				Intelligibility-weighted noise cancellation (dB)	Net intelligibility-weighted gain (dB)
Test	M1 *"armchair"*	M2 *"playground"*	F1 *"pancake"*	F2 *"woodwork"*		
#1	−75°	0°	20°	75°		
	8.04	(0.15)	4.98	3.07	9.25	9.09
#2	30°	−45°	60°	−10°		
	8.34	4.71	4.12	(0.67)	8.38	7.71
#3	10°	−80°	−50°	45°		
	(0.55)	6.90	5.57	3.83	8.56	8.00
#4	−30°	15°	5°	−60°		
	10.53	2.07	(1.14)	6.35	8.27	7.13
#5	−25°	25°	−70°	80°		
	8.09	(0.34)	5.82	4.46	8.78	8.44

TABLE II. Same as Table I except that the recordings were made in a moderately quiet conference room with a 400 ms reverberation time [RT was derived using Schroeder's method; see J. Acoust. Soc. Am. 37, 409–412 (1965)].

Test	Intelligibility-weighted signal cancellation (dB)				Intelligibility-weighted noise cancellation (dB)	Net intelligibility-weighted gain (dB)
	M1 *"armchair"*	M2 *"playground"*	F1 *"pancake"*	F2 *"woodwork"*		
#1	−75°	0°	20°	75°		
	4.82	(0.55)	4.07	2.06	6.73	6.18
#2	30°	−45°	60°	−10°		
	6.27	4.18	3.09	(0.58)	7.26	6.69
#3	10°	−80°	−50°	45°		
	(1.12)	3.85	2.91	2.71	5.75	4.63
#4	−30°	15°	5°	−60°		
	6.29	(0.85)	0.91	3.61	6.16	5.25
#5	−25°	25°	−70°	80°		
	5.70	(0.69)	4.28	2.92	6.97	6.29

while Table II the results from the conference room test. In the tables, the numbers in parentheses represent the degree of cancellation in dB of the desired source (which should ideally be 0 dB) and the other numbers represent the degree of noise cancellation for each noise source. Because we had separate recordings of speech signals from each talker, we applied the same processing both on the complex signal and synchronously on each signal corresponding to each talker as well. As such, we were able to tell the effect of processing on each signal involved. The next to the last column in the tables show the degree of cancellation for all the noise sources lumped together, while the last column gives the net intelligibility-weighted improvement (which considers both noise cancellation and loss in the desired signal). Our results from the anechoic room show that, in the intelligibility-weighted measure, the cancellation strategy was able to cancel each noise source by 3–11 dB, while the degradation in the desired source was very small (mostly smaller than 0.5 dB). The total noise cancellation was between 8 and 10 dB. For the conference room, the cancellation was roughly 2 dB less, indicating that the room reverberation degraded the system performance somewhat. In spite of the drop in system performance the system still produced a sizable gain in speech intelligibility.

In order to get an insight into the effect of the signal processing on each talker, we choose one subtest example (anechoic room; desired source: F1 at 60°; noise sources: M1 at 30°, M2 at −45°, and F2 at −10°). We display the signal waveform of each talker as well as the complex signal of all the four talkers, before [Fig. 7(A)] and after [Fig. 7(B)] the signal processing. Comparison of the two panels shows a great attenuation of the interfering talkers (M1, M2, and F2) while the desired signal (F1) is essentially not attenuated and the distortion of the desired talker is unperceivable. A moment-by-moment comparison shows that the momentarily strongest noise source was always reduced, indicating that the system adapted rapidly. The last trace in Fig. 7(B) is the system output, which turned out to be cleaner and closer to the desired speech [F1 in Fig. 7(A)] than the noisy unprocessed signal [the last trace in Fig. 7(A)].

In an informal listening experiment with normal hearing listeners, we found the unprocessed signal to be impossible to understand, even when the spatial cues were retained. After the processing, however, the extracted speech from a desired source was easily understandable.

Limited by our experimental facility, we only conducted on-site acoustic tests for four-talker situations. However, our computer simulation results for six-talker situations were quite similar. To avoid redundancy, we omit presentation of the details. Basically, we obtained a 7–10 dB enhancement in the intelligibility-weighted signal-to-noise ratio when there were six equally loud, temporally aligned speech sounds originating from six different sources.

V. DISCUSSION

There are three key differences between the algorithm proposed in this paper and conventional adaptive beamformers such as the Frost and Griffiths-Jim beamformers (Van Veen and Buckley, 1988), namely, (i) direct frequency-domain null steering, (ii) explicit source localization, and (iii) implicit utilization of dialogue characteristics. The frequency-domain null-steering algorithm described herein does rapid, independent steering of the beampattern at each frequency. Independent steering allows rapid steering of the single null at each frequency to the dominant interferer at that time and frequency. It provides a maximum potential to cancel intense components emitted from multiple interferers with only two inputs available. What distinguishes this method from other methods is that this independent steering can be implemented with no time delay when it is provided with instant localization information. When processing signals with strong, rapidly varying time-frequency structure such as speech, the net effect is to allow significant cancellation of several *simultaneous* interferers by exploiting differences in their instantaneous time-frequency structures. In contrast, slowly adapting time-domain algorithms such as the Frost (Frost, 1972) and Griffiths-Jim (Griffiths and Jim, 1982) beamformers are unable to track the nonstationary structure rapidly enough to achieve significant cancellation of more than a single interferer. This claim is clearly dem-

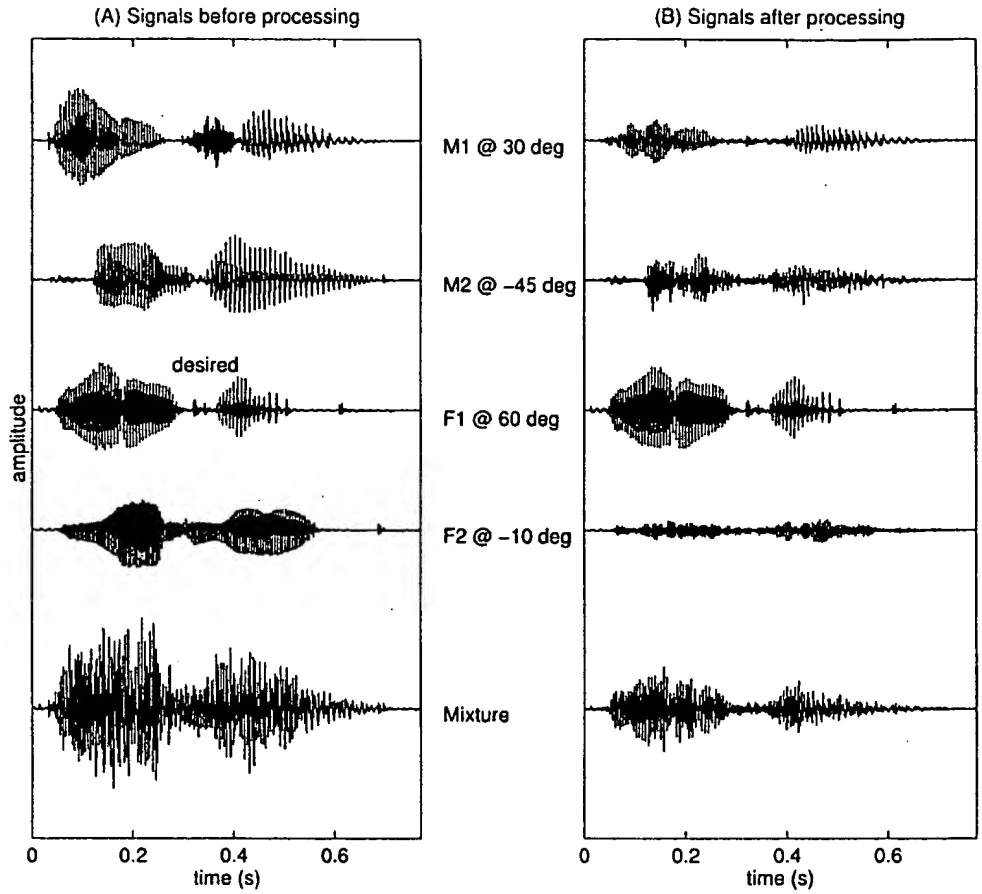


FIG. 7. The signal waveform of each talker as well as the complex signal of all the four talkers before (A) and after (B) the signal processing. See Fig. 6 for the test configuration.

onstrated in the results of comparison experiment presented in Yang *et al.* (2000) using complete sentences under a variety of signal-to-noise conditions.

The performance of this algorithm is comparable to the conventional beamformers for the case of a single interferer, but markedly better for cases involving more than one interferer. The comparisons conducted in Zheng *et al.* (2001) were made in computer simulation where up to four interferers were used at four different SNR settings (−6, −3, 0, +3 dB). In the presence of two or more interferers, the present method provided 6–7 dB of SNR gains, while the Frost beamformer and the Griffiths-Jim beamformer had SNR gains in the 2–5 dB range.

The second difference between the conventional beamformers and the proposed method is that the latter explicitly identifies the spatial directions of the target and interferers via a nonlinear, cross-frequency localization procedure (Liu *et al.*, 2000) and exploits this information to optimally steer the null pattern in each frequency bin. The localization is conducted on a frame-by-frame basis and the results are used immediately by the cancellation on the same frame. Therefore, as mentioned above, the adaptation time is virtually zero. This feature is especially important when processing signals with rapidly varying time-frequency structure such as speech. Explicit source localization also offers several other potential advantages, including the ability to steer toward a spatially moving target, better and more robust estimation of signal and interference locations from which to optimize the beam patterns, and the ability (not explored here) to perform additional useful tasks such as auditory scene characterization. The results in Zheng *et al.* (2001) suggest that these characteristics may indeed be advantageous in many situations (with different number of interferers, different spatial configurations, and different SNRs), particularly when the interferers are in close azimuthal proximity to the target.

The third, and most unique, difference is that our method takes full advantage of the characteristics and masking effect of human dialogue as detailed in Sec. III. That strategy makes it possible to utilize a limited resource (two inputs only) to obtain maximum speech intelligibility enhancement benefits such as effective cancellation of multiple interfering sources.

The improvement in signal quality reported in Tables I and II is encouraging but preliminary. The algorithm's performance in anechoic conditions (8–10 dB cancellation) is sufficient to justify further research, while the performance in the conference room (2 dB less cancellation) raises the question as to whether, when used in a real-time environment, the quality of the cancellation will degrade so as to no longer be useful. Practical computational limitations restricted the work reported here, although improvements have allowed off-line analysis over a wider range of materials (Zheng *et al.*, 2001). A related frequency domain beamformer (Lockwood *et al.*, 1999) has been implemented in

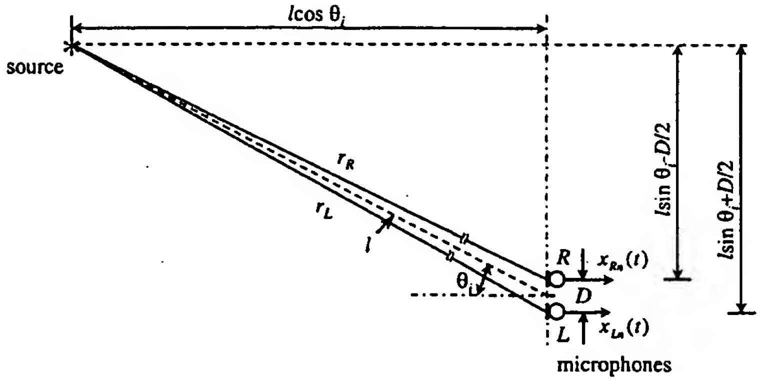


FIG. 8. Top view of the geometry of the source-microphone distance.

real-time (Elledge, 2000) with highly satisfying subjective sound improvement and quality (Larsen *et al.*, 2001). A real-time version of the present algorithm is in the process of implementation; it should permit subjective evaluations to determine whether the technique is viable for hearing aid and other applications.

One practical issue is that when the source to microphone distance is very short (e.g., 2 m or less), it is important to compensate for left–right differences in channel intensity; indeed preliminary tests indicated degradation of about 1 dB in the total net gain without compensation. However, for larger source-microphone distances (e.g., >2 m), the difference with and without compensation was insignificant.

VI. SUMMARY AND CONCLUSION

In this paper, we have presented the technique and experimental results that illustrate the performance of signal processing systems designed for effective extraction of a desired signal in the presence of multiple competing talkers. The signal processing technique is based on dual delay-line structure, a well-known biological network for binaural hearing. The entire system consists of two steps: localization of all sources and extraction of the desired source. Our anechoic chamber tests showed an 8–10 dB of speech enhancement in the presence of four equally loud, temporally aligned talkers; our computer simulation showed a 7–10 dB of speech enhancement in the presence of six equally loud, temporally aligned talkers. The system can localize all the sources present and allow the user to selectively extract any one of them, hence it is more flexible than assuming that the desired source is always straight ahead. It can be applied in many applications such as radar, sonar, communications, and robots.

It is noted that in the present study we focused on separating out a particular talker from all the other competing talkers, i.e., selective hearing. It is technically straightforward to convert the present system to a simulator that can capture the source to which the gaze of the listener is directed at any time instant. It is also possible to simply use multiple noise-cancellation components following the localization so as to extract each of the sources within the environment, i.e., to achieve separation of multiple signals.

The dual delay-line structure implies that the computation is highly parallel. That, and the repeated use of the Fourier transform, made it practical to implement the algorithm by means of VLSI for a fast, miniature device.

Our future work includes evaluation using formal tests in normal listening rooms with human subjects with real-time versions of the algorithm. We will also extend our algorithms to compensate for reverberant environments.

ACKNOWLEDGMENTS

This research was supported by grants from the Beckman Institute, UIUC and from the National Institute for Deafness and other Communication Disorders of the NIH (R21 DC 04870). The work was conducted when the first author (CL) was a Beckman Fellow. A number of people from the Beckman Institute generously donated their time to participate in the recording of speech materials used for our tests. Their generosity is acknowledged. We thank Dean Garstecki for permission to use the anechoic room at the Northwestern University. We also wish to thank the two anonymous reviewers for their constructive comments to an earlier version of the paper.

APPENDIX A: CALCULATION OF THE AMPLITUDE FACTORS

In this appendix we illustrate the calculation of the amplification factors $\alpha_i(m)$, which are used to compensate for differences in the amplitudes of signals arriving at the two microphones. In this example we model the sound source as a simple point source, ignore the absorption of energy by the media, and assume the amplitude variation is solely dependent on the differences in the distance from source to the microphones. In this case, the amplitude compensation is independent of frequency.

The amplitude of the received sound pressure $|\mathbf{p}|$ varies with the source–receiver distance r:

$$|\mathbf{p}| \propto \frac{1}{r} \tag{A1}$$

or

$$\frac{|\mathbf{p}_L|}{|\mathbf{p}_R|} = \frac{r_R}{r_L}, \tag{A2}$$

where $|\mathbf{p}_L|$ and $|\mathbf{p}_R|$ are the amplitude of sound pressures at the two microphones (Kinsler *et al.*, 1982, p. 168). According to the geometry in Fig. 8, the distances from the source

to the left and right microphones r_L and r_R are, respectively,

$$r_L = \sqrt{(l \sin \theta_i + D/2)^2 + (l \cos \theta_i)^2}$$
$$= \sqrt{l^2 + lD \sin \theta_i + D^2/4} \quad \text{(A3)}$$

and

$$r_R = \sqrt{(l \sin \theta_i - D/2)^2 + (l \cos \theta_i)^2}$$
$$= \sqrt{l^2 - lD \sin \theta_i + D^2/4}. \quad \text{(A4)}$$

For a pair of taps in the dual delay-line in Fig. 2, in order to equalize the signals at the tap outputs, the compensation factors $\alpha_i(m)$ and $\alpha_{I-i+1}(m)$ must satisfy the following condition:

$$|\mathbf{p}_L|\alpha_i(m) = |\mathbf{p}_R|\alpha_{I-i+1}(m). \quad \text{(A5)}$$

Substituting Eq. (A2) into Eq. (A5), the above condition becomes

$$\frac{r_L}{r_R} = \frac{\alpha_i(m)}{\alpha_{I-i+1}(m)}. \quad \text{(A6)}$$

We define the value of $\alpha_i(m)$ to be equal to

$$\alpha_i(m) = K\sqrt{l^2 + lD \sin \theta_i + D^2/4}, \quad \text{(A7)}$$

where K has unit of inverse length and is chosen for a convenient amplitude level. Applying the definition in Eq. (A7), the value of $\alpha_{I-i+1}(m)$ will be

$$\alpha_{I-i+1}(m) = K\sqrt{l^2 + lD \sin \theta_{I-i+1} + D^2/4}$$
$$= K\sqrt{l^2 - lD \sin \theta_i + D^2/4}, \quad \text{(A8)}$$

where the relationship $\sin \theta_{I-i+1} = -\sin \theta_i$ can be obtained by substituting $I-i+1$ into i in Eq. (3). By substituting Eqs. (A7) and (A8) into Eq. (A6), one can verify that the values assigned to $\alpha_i(m)$ in Eq. (A7) satisfy the condition in Eq. (A6).

APPENDIX B: DEFINITION OF THE INTELLIGIBILITY-WEIGHTED MEASURE

For any signal s, the intelligibility-weighted measure $\Gamma(s)$ is calculated by (Link and Buckley, 1993)

$$\Gamma(s) = \int_{BW} W_{\mathrm{AI}}(f) 20 \log_{10} \mathrm{rms}_{1/3}(|S(f)|) df, \quad \text{(B1)}$$

where

$$W_{\mathrm{AI}}(f) = \frac{1/[1 + (f/1925)^2]}{\int_{BW} 1/[1 + (f/1925)^2] df}, \quad \text{(B2)}$$

$$\mathrm{rms}_{1/3}(|S(f)|) = \left[\frac{\int_{2^{-1/6}f}^{2^{1/6}f} |S(f')|^2 df'}{(2^{1/6} - 2^{-1/6})f} \right]^{1/2} \quad \text{(B3)}$$

and BW denotes the frequency range. The system improvement in the intelligibility-weighted measure for the target signal T and interference I are, respectively,

$$\Delta\Gamma(T) = \Gamma(T_o) - \Gamma(T_i) \quad \text{(B4)}$$

and

$$\Delta\Gamma(I) = \Gamma(I_i) - \Gamma(I_o), \quad \text{(B5)}$$

where the subscripts i and o denote the input and output, respectively. The overall (or net) intelligibility-weighted gain, G_I, is the sum of the two measures, thus

$$G_I = \Delta\Gamma(T) + \Delta\Gamma(I). \quad \text{(B6)}$$

In our experiment, since we had separate recordings of the target and noise signals, i.e., T_i and I_i (the latter might include more than one interfering talker), we were able to apply the same processing framewise on either of them and obtain the results T_o and I_o, respectively. The framewise processing [Eq. (11)], however, was determined based on the target and noise signals mixed together as would be encountered in the real situation. It is noted that the spectrum $S(f)$ was computed based on the full-length signal, which in our case was a whole spondaic word. We used the full long-term spectrum, as opposed to a frame-by-frame spectrum, for two reasons: (1) It was consistent with the way the intelligibility-weighted measure was applied in other papers published in the area such that our results could be compared directly with earlier results; (2) Since our system has virtually no adaptation time (i.e., it almost always is successful in localizing the strongest interference within a few milliseconds), there is no advantage to computing with the short-term spectra.

Since the intelligibility-weighted measure was constructed as an estimate of the subjective improvement based on the objective calculation, it deliberately emphasized the low frequency domain according to the "critical-band" theory. However, because the low frequency domain is always the hardest to clean with the approaches using multimicrophone arrays of limited size, the intelligibility-weighted measure usually has a smaller value (~1 dB difference in our experiment) than the non-weighting counterpart, i.e., SNR. Nonetheless, this effect does not change the overall picture of the performance; especially the comparison of our system with others, as given in the paper, remains valid.

Similarly, if we denote the array beampattern as $E(f, \theta)$, where f is frequency and θ is the incident direction, the intelligibility-weighted beampattern, $\bar{E}(\theta)$, is defined by (Liu and Sideman, 1996)

$$\bar{E}(\theta) = \int_{BW} W_{\mathrm{AI}}(f) E(f, \theta) df. \quad \text{(B7)}$$

Banks, D. (1993). "Localization and separation of simultaneous voices with two microphones," IEE Proc. I **140**, 229–234.

Blauert, J. (1983). *Spatial Hearing: The Psychophysics of Human Sound Localization*, John S. Allen, translator (The MIT Press, Cambridge, MA).

Bodden, M. (1993). "Modeling human sound source localization and the cocktail-party-effect," Acta Acust. **1**, 43–55.

Bodden, M. (1996). "Auditory demonstration of a cocktail-party-processor," Acust. Acta Acust. **82**, 356–357.

Bregman, A. S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound* (The MIT Press, Cambridge, MA).

Bronkhorst, A. W., and Plomp, R. (1992). "Effect of multiple speechlike maskers on binaural speech recognition in normal and impaired hearing," J. Acoust. Soc. Am. **92**, 3132–3139.

Chabries, D. M., Christiansen, R. W., Brey, R., and Robinette, M. (1982). "Application of the LMS adaptive filter to improve speech communication in the presence of noise," Proc. IEEE ICASSP, 148–151.

Chazan, D., Medan, Y., and Shvadron, U. (1988). "Noise cancellation for hearing aids," IEEE Trans. Acoust., Speech, Signal Process. 36, 1697–1705.

Cherry, E. C. (1953). "Some experiments on the recognition of speech, with one and with two ears," J. Acoust. Soc. Am. 25, 975–979.

Colburn, H. S., and Durlach, N. I. (1978). "Models of binaural interaction," in *Handbook of Perception, IV: Hearing*, edited by E. C. Carterette and P. F. Morton (Academic, New York).

Colburn, H. S. (1996). "Computational models of binaural processing," in *Auditory Computation*, edited by H. L. Hawkins, T. A. McMullen, A. N. Popper, and R. R. Fay (Springer, New York), pp. 332–400.

Durlach, N. I. (1960). "Note on the equalization and cancellation theory of binaural masking level differences," J. Acoust. Soc. Am. 32, 1075–1076.

Durlach, N. I. (1972). "Binaural signal detection: Equalization and cancellation theory," in *Foundations of Modern Auditory Theory*, edited by V. T. Jerry (Academic, New York), pp. 369–462.

Elledge, M. E., Lockwood, M. E., Bilger, R. C., Feng, A. S., Jones, D. L., Lansing, C. R., O'Brien, W. D., and Wheeler, B. C. (2000). "Real-time implementation of a frequency-domain beamformer on the TI C62X EVM," the 10th Ann. DSP Technol. Educ. Res. Conf., Houston, TX, August 2–4, 2000.

Feng, A. S., and Shofner, W. P. (1981). "Peripheral basis of sound localization in anurans: Acoustic properties of the frog's ear," Hear. Res. 5, 201–216.

Flanagan, J. L. (1972). *Speech Analysis, Synthesis and Perception* (Springer-Verlag, Berlin).

Frost, O. L. (1972). "An algorithm for linearly constrained adaptive array processing," Proc. IEEE 60, 926–935.

Gaik, W. (1993). "Combined evaluation of interaural time and intensity differences: Psychoacoustic results and computer modeling," J. Acoust. Soc. Am. 94, 98–110.

Griffiths, L. J., and Jim, C. W. (1982). "An alternative approach to linearly constrained adaptive beamforming," IEEE Trans. Antennas Propag. AP-30, 27–34.

Jeffress, L. A. (1948). "A place theory of sound localization," J. Comp. Physiol. Psychol. 41, 35–39.

Jeffress, L. A. (1970). "Masking," in *Foundations of Modern Auditory Theory*, edited by J. V. Tobias (Academic, New York), pp. 85–114.

Kinsler, L. E., Frey, A. R., Coppens, A. B., and Sanders, J. V. (1982). *Fundamentals of Acoustics* (Wiley, New York).

Konishi, M., Takahashi, T., Wagner, H., Sullivan, W. E., and Carr, C. E. (1988). "Neurophysiological and anatomical substrates of sound localization in the owl," in *Auditory Function: Neurobiological Bases of Hearing*, edited by M. E. Gerald, W. E. Gall, and W. M. Cowan (Wiley, New York), pp. 721–745.

Larsen, J., Lockwood, M. E., Lansing, C. R., Bilger, R. C., O'Biren, W. D., Jones, D. L., Wheeler, B. C., and Feng, A. S. (2001). "Performance of a frequency-based minimum variance beamforming algorithm for normal and hearing impaired listeners," J. Acoust. Soc. Am. 109, 2494.

Lindemann, W. (1986). "Extension of a binaural cross-correlation model by contralateral inhibition. I. Simulation of lateralization for stationary signals," J. Acoust. Soc. Am. 80, 1608–1622.

Link, M. J., and Buckley, K. M. (1993). "Prewhitening for intelligibility gain in hearing aid arrays," J. Acoust. Soc. Am. 93, 2139–2145.

Liu, C., and Sideman, S. (1996). "Simulation of fixed microphone arrays for directional hearing aids," J. Acoust. Soc. Am. 100, 848–856.

Liu, C., Wheeler, B. C., O'Brien, Jr., W. D., Bilger, R. C., Lansing, C. R., and Feng, A. S. (2000). "Localization of multiple sound sources with two microphones," J. Acoust. Soc. Am. 108, 1888–1905.

Lockwood, M. E., Jones, D. L., Elledge, M. E., Bilger, R. C., Feng, A. S., Goueygou, M., Lansing, C. R., Liu, C., O'Brien, Jr., W. D., an Wheeler, B. C. (1999). "A minimum variance frequency-domain algorithm for binaural hearing aid processing," J. Acoust. Soc. Am. 106, 2278.

Peterson, P. M. (1989). "Adaptive array processing for multiple microphone hearing aids," Ph.D. Dissertation, Dept. Elect. Eng. and Comp. Sci., MIT; Res. Lab. Elect. Tech. Rept. 541, MIT, Cambridge, MA.

Peterson, P. M., Durlach, N. I., Rabinowitz, W. M., and Zurek, P. M. (1987). "Multimicrophone adaptive beamforming for interference reduction in hearing aids," J. Rehabil. Res. Dev. 24, 103–110.

Stern, R. M., and Trahiotis, C. (1995). "Models of binaural interaction," in *Hearing*, edited by B. C. J. Moore (Academic, San Diego), pp. 347–386.

Stern, R. M., and Trahiotis, C. (1997). "Models of binaural perception," in *Binaural and Spatial Hearing in Real and Virtual Environments*, edited by R. H. Gilkey and T. R. Anderson (Lawrence Erlbaum Associates, Mahwah, NJ), pp. 499–532.

Strube, H. W. (1981). "Separation of several speakers recorded by two microphones (cocktail-party processing)," Signal Process. 3, 355–364.

Takahashi, T. T., and Keller, C. H. (1994). "Representation of multiple sound sources in the owl's auditory space map," J. Neurosci. 14, 4780–4793.

Van Veen, B. D., and Buckley, K. M. (1988). "Beamforming: a versatile approach to spatial filtering," IEEE ASSP Mag., April 1988, 4–24.

Weiss, M. (1987). "Use of an adaptive noise canceler as an input preprocessor for a hearing aid," J. Rehabil. Res. Dev. 24, 93–102.

Widrow, B., Glover, J. R., McCool, J. M., Kaunitz, J., Williams, C. S., Hearn, R. H., Zeidler, J. R., Dong, E., and Goodlin, R. C. (1975). "Adaptive noise cancelling: Principles and applications," Proc. IEEE 63, 1692–1716.

Yang, K. L., Lockwood, K. L., Elledge, M. E., and Jones, D. L. (2000). "A comparison of beamforming algorithms for binaural acoustic processing," Proceedings of the 9th IEEE Digital Signal Processing Workshop, Hunt, TX, October 15–18, 2000.

Yin, T. C. T., and Chan, J. C. K. (1990). "Interaural time sensitivity in medial superior olive of cat," J. Neurophysiol. 64, 465–488.

Zheng, Y., Lockwood, M. E., Wheeler, B. C., Jones, D. L., Feng, A. S., O'Brien, W. D., Bilger, R. C., and Lansing, C. R. (2001). "Comparison of binaural beamformers for speech extraction in complex auditory scenes," J. Acoust. Soc. Am. 109, 2494.